TI 2023-058/I
Tinbergen Institute Discussion Paper

Why do oaths work?
Image concerns and credibility in promise keeping

Sorravich Kingsuwankul[1]
Chloe Tergiman[2]
Marie Claire Villeval[3]

1 Vrije Universiteit Amsterdam and Tinbergen Institute
2 Pennsylvania State University
3 Univ Lyon

Why do oaths work?
Image concerns and credibility in promise keeping*

Sorravich Kingsuwankul† Chloe Tergiman‡ Marie Claire Villeval§

September 16, 2023

Abstract

We use a laboratory experiment to understand the channels through which honesty oaths can affect behavior and credibility. Using a game with asymmetric information in a financial market setting that captures some important features of advisor-investor interactions, we manipulate the common knowledge of the promise and investigate three non-pecuniary costs of breaking an oath: co-player image costs, audience-image costs, and self-image costs. For investors oaths are neither sufficient nor necessary to generate trust: ultimately investors rely on their experience. We link laboratory results to a survey we conducted in the Netherlands where oaths are required in the banking sector.

Keywords: Promise-keeping, Honesty Oaths, Common Knowledge, Deniability, Financial Markets, Laboratory Experiment

JEL codes: C91, D01, D83, D91

*Acknowledgements: We are grateful to B. Corgnet, L. Gangadharan, F. Gul, N. Jacquemet, R. Kerschbamer, C. Raymond, H. Schneider, V. Valero, and participants at presentations at the Applied Ethic and Morality Research Group at the University of Prague, the Credence Goods and Expert Markets Webinar, Cornell University, Michigan Ross TO, Penn State University, Stanford University, Queen's University, the East Coast Conference, ASFEE in Montpellier, MBEES-MBEPS in Maastricht, the ESA World meeting in Lyon, the TIBER symposium in Tilburg, and the European Economic Association meeting in Barcelona for useful comments. We are thankful to Q. Thevenet for his assistance in programming the experiment. This research has benefited from the financial support of IDEXLYON from Université de Lyon (INDEPTH-IDEX/SBP/2018/03) within the Programme Investissements d'Avenir (ANR-16-IDEX-0005) operated by the French National Research Agency, and of the LABEX CORTEX (ANR-11-LABX-0042) of the Université de Lyon, within the Programme Investissements d'Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR).

†School of Business and Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Tinbergen Institute, 1082 MS Amsterdam, The Netherlands. Email: s.kingsuwankul@vu.nl

‡Pennsylvania State University, Smeal College of Business, 334 Business Building, University Park, PA 16802, United States. Email: cjt16@psu.edu

§Univ Lyon, CNRS, GATE UMR 5824. 93, Chemin des Mouilles, F-69130 Ecully, France. IZA, Bonn, Germany. Email: villeval@gate.cnrs.fr

1 Introduction

A number of well publicized frauds (*e.g.*, the Madoff, Enron, Wells Fargo and the Libor rate-fixing scandals) have decreased the public's trust towards the financial sector (Guiso et al., 2008; Sapienza and Zingales, 2012). In fact, "trust issues" are endemic to financial markets, leading individuals to exit these markets or to not participate in them altogether, forgoing possibly lucrative opportunities.[1] Instituting a rule that a banker must take an oath has been proposed as a tool to reduce fraud and restore public trust. Analogous to the Hippocratic Oath for medical doctors, a banker oath aims to commit employees in financial institutions to prescribed ethical standards, maintain transparency and accountability, and place a client's interest before personal ones (Boatright, 2013). The Netherlands was the first country to implement such a practice. Since 2015 all employees in the Dutch banking sector are legally obliged to swear this oath prior to starting their jobs.[2] A similar intervention was adopted in Belgium in 2019. However, such country-wide oaths remain uncommon.[3]

The main goal of this study is to evaluate the channels through which an oath can have an impact. Towards this end, and with an aim to keep close to a financial setting, we designed and experimentally tested a game with asymmetric information where an advisor is aware of how likely an investment is to yield a return, but an investor is not. We focus first on the behavior of the oath-taker: if an oath works, why does it work? We draw from the literature on preferences for promise-keeping and truth-telling (reviewed in the next section), and investigate the importance of three non-pecuniary costs of breaking a promise: co-player-image costs, audience-image costs, and self-image costs. We define co-player-image costs as those that a player can incur because of how other players in the game may perceive them if they appear to have broken a promise.[4] Audience-image costs are generated by audience members, which may include the experimenter or God.[5] Finally, self-image costs are those

[1] Johnson et al. (2016) found that a reason behind a sluggish mortgage re-financing was a feeling of suspicion — borrowers perceived the offer "too good to be true", anticipated hidden costs and, therefore, simply left money on the table.

[2] Loonen and Rutgers (2017) provides a summary on the introduction of bankers' oath in the Netherlands. See also: https://www.tuchtrechtbanken.nl/en/about-the-bankers-oath/, accessed on July 17, 2023.

[3] Professional oaths are not a recent phenomenon, as some date back to the Middle Ages (for example, in craft guilds, https://www.sacred-texts.com/mas/dun/dun03.htm, accessed on July 17, 2023.). In addition, if corporate codes of ethics become widespread (*e.g.*, Babri et al., 2021), asking professionals to swear an oath remains limited to a few professions in order for a licence to be issued. For example, the Hippocratic oath, which can be traced to the 5th century, is an important step to become a doctor. Lawyers and accountants also have to swear an oath to be able to practice (see Douglass, 1995; Andrews, 2009; Heese et al., 2023).

[4] These costs may prevent an individual from breaking a promise if that individual does not want to be perceived as someone who would indeed not keep their word (see, *e.g.*, Dufwenberg and Dufwenberg, 2018; Gneezy et al., 2018; Abeler et al., 2019) or does not want to be seen as someone for whom it is easy to break a promise, that is, someone with a low private cost of doing so (see, *e.g.*, Frankel and Kartik, 2019).

[5] The audience differs from other players in that the audience makes no decisions in the game but may hold beliefs about a player after observing their actions. In thinking about the audience we thus share the view of Abeler et al. (2019).

costs that may prevent an individual from breaking a promise even if no one else can observe it. They represent the disutility from seeing oneself as having a low moral identity.[6]

Mirroring our first question (the impact of taking an oath on the oath-taker), we also study how oaths impact those who rely on the oath-takers's actions, which in our game are the investors: is knowledge of the oath necessary or sufficient to increase the oath-taker's credibility? That is, we investigate whether investors need to be aware of the oath for it to have an impact on their behavior toward the oath-taker, or whether they instead predominantly rely on their experience, independently of the knowledge that an oath has been taken.

In our game, a financial advisor receives three cards at random, where each card has a 0.5 probability to display a star and otherwise is blank. The probability that an investment with the financial advisor leads to a successful return for an investor is determined by the fraction of cards that have a star on them. Thus, the cards can be viewed as representing the quality of the advisor's portfolio. Under asymmetric information, the advisor, who can see the face of all three cards, sends a cheap-talk message about the number of stars to a player in the role of an investor who cannot see the face of the cards.[7] The investor then decides whether to invest with the advisor or not. Finally, Nature draws one of the advisor's three cards to determine payment for the investor: the investment is a success if the drawn card has a star on it, and is a failure if the drawn card is blank. The advisor's payment, however, only depends on whether an investment was made or not, and not on the success of the investment, creating a conflict of interest between the two players when the state of the world is bad.[8]

As is the case in life outside the laboratory, and in the financial sector in particular, in our setting lies can be *detectable* or *deniable*. The former happen when the advisor announces three stars after observing fewer than three stars — in these situations, the advisor's lie is exposed if Nature draws a blank card. The latter happen when, for example, the advisor

[6]These costs may result from the fact that an oath encourages moral deliberation in oath-takers (de Bruin, 2016) or from an intrinsic preference for keeping one's words resulting from the internalization of a moral norm (Ellingsen and Johannesson, 2004). One may also interpret self-image costs as costs that come from being one's own audience, as in Dufwenberg and Dufwenberg (2018).

[7]In this setting asymmetric information concerns the characteristics of the good, not the fit between the product and the needs of the customer, which is another source of concern in credence goods markets.

[8]Evidence in the field shows that the incentive scheme often rewards advisors based on the sales of financial products and can lead them to give biased recommendations to their clients (see Mullainathan et al., 2012; Anagol et al., 2017; Pool et al., 2016; Foerster et al., 2017). This gives rise to self-serving tendencies where advice may not be in the best interest of the clients. A study has shown that the bonus share in bank employees' compensation correlates with an increase in the relative trading income of the bank (Efing et al., 2015). Financial advisors' bonuses do not always depend on the success of these products. Salesforces may be biased also because they receive commissions from the providers of financial products (Inderst and Ottaviani (2012b); see Gravelle (1993); Inderst and Ottaviani (2012a) for models), or from trading bank proprietary portfolios that tend to pay lower returns to retail investors on average (Fecht et al., 2018).

announces two stars after observing fewer than two stars: in these situations, the advisor can maintain plausible deniability regardless of Nature's draw since whatever the draw, it will not contradict the two-star announcement.[9] Allowing for a richer set of lies not only makes our setting more realistic, but also allows us to more precisely understand the mechanism through which an oath can impact behavior.

Our five between-subject treatments, run with 607 participants, re-match participants across periods and vary whether and how an oath is introduced.[10] In the ***No-Oath*** treatment, advisors do not swear any oath. This gives us a baseline for the distributions of the announcements by the advisors and of the investment rates for each type of announcement by the investors. In the ***Oath Common Knowledge (Oath-CK)*** treatment, advisors swear a compulsory oath to be truthful, and this is common knowledge among all participants. Comparing data from the No-Oath and Oath Common Knowledge treatments allows us to directly measure the impact of such an oath on behavior. Overall, the fraction of fully truthful advisors goes from about 10% in the absence of an oath, to about 75% when a common-knowledge oath is present.Common-knowledge oath-taking has a dramatic impact on all types of lies: the fraction of advisors who make detectable lies drops significantly from 71% to 12%, and the fraction who engage in deniable lies goes from 87% to 23%, a drop that is even more surprising since by definition these lies cannot be detected by investors.

After documenting the dramatic impact of the oath on advisors' behavior, we investigate why the oath is so effective. When someone promises to tell the truth, which non-pecuniary costs explain promise-keeping: are individuals primarily concerned about how they are perceived by others, whether it be by another player (co-player-image costs) or the audience (audience-image costs)? Or do they also care about about their own self-image?

To evaluate the impact of co-player-image costs, we designed the ***Oath Private Knowledge (Oath-Private)*** treatment. In this treatment, advisors swear an oath, but investors are kept in the dark regarding its existence, a fact that the advisors are aware of. Comparing behavior in the Oath-CK and Oath-Private treatments allows us to explore the impact of co-player-image costs. Indeed, in both treatments the audience knows the truth regarding the number of stars, as well as the annoucement made. The only difference between these treatments is whether the investor is aware of the oath, that is, whether co-players-image costs can exist vis-a-vis an advisor who breaks a promise. When the oath is privately taken,

[9]This game, known as the "Announcement Game," was first described in Tergiman and Villeval (2023) and stands out from others in that types of lies available (detectable or deniable) are not only richer than in other works, but also the degree of detectability is endogenously chosen by the announcer, features that are closer to those in real financial settings.

[10]The stranger matching protocol allows us to focus on the impact of oaths rather than on other elements that would arise if participants were in fixed pairs or if the history of play was known. In particular, this eliminates the possibility of direct punishment that would exist in case of fixed matching.

"only" about 40% of advisors consistently tell the truth. Therefore, co-player-image costs can account for the additional 35 percentage point increase in truth-tellers that we observe in the Oath-CK treatment where about 75% of advisors tell the truth at all times.

To evaluate self-image and audience-image costs we compare the No-Oath and Oath-Private treatments. Indeed, both these treatments keep co-player-image concerns toward investors constant: investors face identical instructions across these two treatments and, thus, have no reason to view advisors differently in the two treatments given a particular announcement and a particular card drawn by Nature. In other words, because signing the oath is a private event, it cannot lead to additional image costs induced by what investors would think about an advisor who would break that oath. The two treatments instead differ in the following way: while in both treatments an advisor who announces a number of stars that differs from the truth may incur a non-pecuniary cost of lying about the true number of stars, advisors who lie in the Oath-Private treatment may incur additional costs: the self-image and audience-image costs associated with breaking a promise. Since in the No-Oath treatment the fraction of truth-tellers is about 10% and is about 40% in the Oath-Private treatment, we find that self- and audience-image concerns can jointly contribute to 30% additional subjects moving to tell the truth. Using responses from a post-experimental questionnaire in the Oath-Private treatment, we find that for the modal truth-teller, self-image costs seem to be more important than the audience-image costs.

Using data on oath-breakers, we are further able to shed light on what kinds of modeling choices for the self-image costs are consistent with the data. We find that breaking the oath even when it is private information generates a cost that is consistent with not wanting to lie maximally after promising not to lie at all. This type of cost differs from a simple fixed cost that would be independent of the kind of lie that is made. We also show that this cost differs from a cost that would depend on the size of the lie and that has been studied in the absence of the oath (see Gneezy et al., 2018). Regarding co-player-image costs, our data suggest that they are consistent with caring about not wanting to be seen as someone for whom it is easy to break an oath (judgment on the person) as opposed to not wanting to be seen as someone who has broken an oath (judgment on the action).

Turning to the impact of oaths on investors, we find that the common-knowledge oath leads to higher earnings for investors and that investors interpret announcements more credibly than when there is no oath. The next two treatments aim to understand what role the oath *per se* plays in those results: are investors anticipating more honest behavior from advisors because they took an oath? Or are they reacting to the truthfulness of announcements independently of the oath, that is, are they learning from experience? To answer these questions we conducted the ***Asymmetric-Asynchronous (Asym-Async)*** treatment. In this treatment, all participants are investors who are paired with past advisors from the Oath-CK treatment. They are given "an excerpt of the instructions" from the Oath-CK treatment, which are those very instructions except for the portion regarding the oath. Thus, investors in the Asym-Async treatment face the announcements from past advisors who had taken the common-knowledge oath but they are unaware of the existence of the oath. The investors from the Oath-CK and Asym-Async treatments are thus asymmetrically informed about the oath but face the same set of advisors who, as described above, were largely truthful. Despite this asymmetry, by the end of the sessions we find little difference in average investor behavior in the two treatments. Therefore, the oath *per se* is not necessary to induce trust: truthful announcements lead to trust in these announcements even in a game in which deniable lies are easy to make and at the individual level difficult to detect.

If an oath is not necessary to generate trust, is it sufficient? To answer this question, we designed the ***Symmetric-Asynchronous (Sym-Async)*** treatment. In this treatment, all participants are investors who are given the *full* instructions of the Oath-CK treatment, including the section about the oath. They are then told that in each period they will be randomly matched with a past advisor chosen from a *subset* of all the past advisors from that treatment. We built that subset so that the announcement behavior of the advisors does not differ significantly from that of the advisors in the No-Oath treatment who, as described above, were largely untruthful. Comparing investments between the Sym-Async treatment and the Oath-CK and No-Oath treatments informs us of how much importance knowledge of the oath has on investors and whether this impact is persistent. We find that the investment behavior in the Sym-Async treatment resembles that in the No-Oath treatment: knowledge of the oath is not a sufficient condition for trust to be established.

Finally, we tie our laboratory results with the responses of an online survey we conducted via Prolific with 198 respondents residing in the Netherlands, where bank employees have to take a compulsory oath. Our main questions of interest are the extent of public awareness about the oath, trust towards the banking sector, and whether we can link one to the other. Almost half of the respondents either did not know that bankers' oath existed or thought it was unlikely that such an oath existed. Awareness was equally limited among respondents who had a financial advisor. Thus, even in a country where such an oath is a legal obligation since 2015, the awareness is very far from being universal. Further, while 69% of respondents believe that in general "people can be trusted," this trust does not carry over to bankers, where the level of trust is closer to 18%. Nevertheless, respondents who are aware of the oath are more likely to hold a higher level of trust in bankers. Finally, our survey highlights that the majority of our respondents believe that requiring an oath can be beneficial.

The remainder of this paper is as follows. Section 2 reviews the related literature. Section

3 outlines the experimental design and procedures. Section 4 reports the results of the experiment and of the survey. Section 5 discusses these findings and concludes.

2 Related literature

Our study contributes first to the literature on oath-taking. Table A1 in Online Appendix A provides an overview of the experimental literature on the effect of oath taking. Among studies that consider strategic settings (top panel), the closest to ours are three papers that study a deception game (Jacquemet et al., 2019; Beck, 2021; Jacquemet et al., 2021), which show that the presence of an oath can lead to a reduction of lying, though the extent of this reduction depends on whether incentives are aligned or not and how the oath is implemented.[11] Their interpretation is that the oath raises moral awareness.

The literature on oaths in the financial sector specifically is remarkably thin. Weitzel and Kirchler (2023) report an audit study on the impact of *reminding* Dutch bankers of their oaths. Being reminded of the oath makes advisors less likely to prioritize the bank's interests (*i.e.,* pushing unnecessary loans) compared to a baseline without reminder. Limited recall of the oath may play a role, but so might moral reminders (bankers had to explain the purpose of the oath), and a fear of punishment (customers may be more likely to file misconduct complaints). Relatedly, Heese et al. (2023) exploit the introduction of an integrity pledge for Dutch professional accounting degree holders and document a reduction in discretionary accruals in firms where the Chief Financial Officers took the pledge, compared to firms with CFOs not subject to the pledge.[12]

While we begin our own analyses by documenting the impact of an oath on lying and on investors' decisions in a financial setting, which in itself differs from the above work, our main objective goes beyond that: it is to understand why an oath can be impactful. Separating the importance of internal and external image costs contributes to clarifying where and for whom an oath can have an impact, and why this impact may not systematically carry-through to individual decision-making tasks where image concerns may not be as salient.

The Announcement Game incorporates dimensions that are important to understand the impact of oaths. In this game, senders endogenously choose what kinds of lies to make - either detectable or deniable, and the receiver can under certain circumstances learn whether

[11] The bottom panel of Table A1 shows that there is less consensus on the impact of oaths in individual decision-making tasks: some find that oaths can be effective (Beck et al., 2020; Jacquemet et al., 2021; Peer and Feldman, 2021), whereas others either find them ineffective (Koessler et al., 2019; Prima et al., 2020; Cagala et al., 2023) or even counter-effective (Cagala et al., 2021). The meta-analysis by Zickfeld et al. (2022) reported an overall negative impact of oaths on lying with some heterogeneity.

[12] Besides oaths and pledges, ethics training of investment advisors has been shown to reduce the likelihood of engaging in misconduct because of a change in normative beliefs (Kowaleski et al., 2020).

a lie occurred but never what the true state of the world was. These features are relevant in studying financial markets where the types of lies that can be made are richer than what protocols implemented in the previous papers have allowed for. The previous studies on oaths in strategic settings have used deception games where the receiver can never detect a lie.[13] In contrast, the Announcement Game allows us to compare the impact of oaths on detectable and deniable lies in the same game, which, we argue, can better inform on the effectiveness of an oath in practice. If an oath were only to be primarily effective when lies are detectable, what use would it have in a world in which lies can also be deniable?

Our study therefore also contributes to the litterature on preferences for promise-keeping, which has used pre-play communication to understand the impact of promises in social dilemma and trust games. Typically in these games a player promises to take a given action if the first mover selects an option that may pay more to both players but that makes him or her vulnerable to the risk of earning less if the promise is broken. This literature shows that even under stranger matching protocols, individuals tend to keep their promises (Ellingsen and Johannesson, 2004; Charness and Dufwenberg, 2006; Serra Garcia et al., 2013; Ismayilov and Potters, 2016; Ederer and Stremitzer, 2017; Casella et al., 2018; Di Bartolomeo et al., 2019) and the effect persists over time (Ederer and Schneider, 2022). Vanberg (2008) demonstrates that promise-keeping is more likely due to an intrinsic preference for keeping one's words – as advanced by Ellingsen and Johannesson (2004)– and not to the anticipation of the guilt experienced from letting down other players' expectations - as advanced by Charness and Dufwenberg (2006). However, the debate is not closed (Ederer and Stremitzer, 2017). Ismayilov and Potters (2016) argue that promise-keeping is associated with a desire for internal consistency rather than a social obligation since promises create commitment even in settings where the promise is not known to the other party.

We differ from these studies in that subjects in our experiment sign a compulsory honesty oath, and do not promise a particular behavior conditional on the choice of a given action by another player; it is a pledge regarding the truthfulness of the exogenous information provided to another player before she takes action. We exclude any instrumental individual reputational concerns since the past oath-breaking by an advisor cannot be traced and punished by investors.[14] Finally, we go deeper into the investigation of the private costs of not keeping one's words by studying if advisors refrain from lying even when investors

[13]In the studies on lying under oath in individual decision-making tasks, lying is either directly detectable (Koessler et al., 2019; Jacquemet et al., 2020; Prima et al., 2020; Peer and Feldman, 2021; Cagala et al., 2023) or deniable - in the sense that they are only detectable statistically (Heinicke et al., 2019; Cagala et al., 2021; Jacquemet et al., 2021; Babin et al., 2022), but not both as in our game.

[14]Regarding how the audience may play a role when promises are made, our study relates to Lang and Schudy (2020) who study political campaign promises in the lab. Their results suggest that the *size* of the audience may not matter: including an additional audience member who is a subject participant in the room does not impact behavior relative to when the only audience would be the experimenter.

are not aware that the advisors took an oath. Note also that in contrast to our study, this literature does not consider that an external audience (such as the experimenter or God) can induce additional moral costs in case of breaking a promise.

The literature on lying (in the absence of promises) has explored the role of intrinsic preferences for truth-telling and the role that reputational concerns may play. Several models converge to identify three main motives that impact behavior. Based on a meta-analysis of studies using the die-under-the-cup paradigm (Fischbacher and Föllmi-Heusi, 2013) or its variants, Abeler et al. (2019) confront different models and show that one in which the decision to lie depends on the monetary benefits of lying, the intrinsic preference for truth-telling in the absence of an audience, and the reputational costs of being perceived as a liar by an outside audience can best fit the data. Inequality aversion, guilt aversion or conformity cannot rationalize the data. Gneezy et al. (2018), Dufwenberg and Dufwenberg (2018), and Khalmetski and Sliwka (2019) also insist on the importance of image concerns, in addition to a fixed direct cost of lying.[15] We analyze advisor behavior in light of these models in a strategic context. We study the role that various image costs of lying under oath play in the decision to break an oath or not. In our strategic setting, when an oath is present, lying costs associated with misrepresenting the number of stars are compounded by the costs of breaking a promise to be honest and thus the moral costs may be more salient.

3 Experimental design and procedures

3.1 Design

The game: To study the impact of oaths on truth-telling and the way people lie in financial markets, we use the Announcement Game of Tergiman and Villeval (2023). In this game, participants are either assigned the role of advisor or investor. In each period, the advisor and the investor start with an endowment of 30 and 100 tokens, respectively. At the beginning of the period, a random draw determines a set of three cards for each advisor, each of which has an independent 0.5 probability to display a star, which indicates a successful project. This probability is known to both the advisor and the investor but the number of stars (0, 1, 2 or 3) is private information to the advisor. After observing the number of stars, the advisor sends a cheap-talk message to the investor regarding her number of stars and, thus, can misreport. The investor then decides whether or not to invest their endowment with the advisor. Next, Nature selects one of advisor's three cards to determine whether the

[15] For Gneezy et al. (2018) the reputational cost depends on the probability that others perceive the individual as a liar. For Khalmetski and Sliwka (2019) an increase in image concerns increases the range of lies. For Dufwenberg and Dufwenberg (2018) belief-dependent lying costs are proportional to the size of the lies perceived by the audience.

investment is a success (if the drawn card displays a star) or a failure (if it is blank).

Irrespective of the investor's decision, the advisor and the investor learn about the outcome of Nature's draw (but we never inform the investor about the content of the three cards). The payoff of the advisor depends on whether the investor invested or not: an advisor earns 230 tokens if the investor invested, 30 otherwise. The payoff of the investor depends on the decision to invest and the outcome of Nature's draw. If an investor decides not to invest, they keep their 100 tokens. If they decide to invest and a star is drawn, the investment is successful and they earn 300 tokens (100-100+300). If instead a blank card is drawn, the investment fails and they earn 30 tokens (100-100+30).

This finitely-repeated game provides a setting in which lying is common-place: not only is truth-telling not supported in equilibrium, but in addition, in our implementation we use a stranger matching protocol through which the advisor and the investor are randomly rematched to form a new pair after each period.[16] As a result we are able to study in what way and how oaths can move participants to be more truthful.

Roles were fixed for the whole session and to avoid end game effects participants were only informed that they would play a minimum of 10 periods and a maximum of 30 periods, though the number of periods, 18, was pre-determined before the experiment. At the end of the session, the program randomly selected one period to count for payment for the Announcement Game.

In part 1 of the experiment, that is before receiving the instructions for the Announcement Game (part 2), participants played eight practice periods of the "Truthful Announcement Game" where they all held the role of investor, and where each participant was paired with a truth-telling computer. These periods aimed to help participants familiarize themselves with the setting and encounter all possible cases of the announcements corresponding to the probabilities used in the Announcement Game. After the eight practice periods, participants made five incentivized decisions: in period 9, participants decided whether or not to invest in each of the four possible scenarios (*i.e.,* 0, 1, 2, 3 stars), and in period 10, they made an investment decision without being informed of the number of stars. Investment decisions in periods 9 and 10 allow us to establish a benchmark of risk attitudes to study investors' behavior in the Announcement Game. Either period 9 or 10 was randomly selected to count for payment for the Truthful Announcement Game, without feedback until the end of the session. Below we describe our five between-subject treatments.[17]

[16]The reason for this choice is that Tergiman and Villeval (2023) showed that this configuration is the one that suffers most from dishonest behavior, that is, both detectable and deniable lies. The impact of oaths that we find in our experiment therefore likely represents a lower bound relative to what may be achieved when players are in fixed partnerships.

[17]The full set of translated instructions are in Online Appendix B.

No-Oath treatment: In the No-Oath treatment, no oaths were taken by participants in the role of financial advisors and subjects played 18 periods of the Announcement Game as it is described above.

Oath Common Knowledge ("Oath-CK") treatment: In the Oath-CK treatment, participants in the role of the advisor were required to swear the following oath: "I swear upon my honor that during this experience I will behave honestly and I will always tell the truth."[18] The instructions for all participants stated that those who would be assigned the role of advisor would be asked to take an oath by which they would swear upon their honor that during the experiment they would behave honestly and always tell the truth. The oath-taking stage took place after participants discovered their role in the game. Advisors swore the oath by typing in the text of the oath on their computers to facilitate its encoding in memory (see Naka and Naoi, 1995; Skinner et al., 1997).[19] The text for the oath and taking of the oath was common knowledge. This treatment allows us to measure the overall impact of the oath on advisors' likelihood of lying and on the frequency of detectable and deniable lies, as well as on investor behavior.

Oath Private Knowledge ("Oath-Private") treatment: In the Oath-Private treatment, the instructions handed out to all participants were identical to those of the No-Oath treatment. However, after discovering their role in the Announcement Game, all the participants in the role of advisor were asked to take the oath, while those in the role of investor, who were not informed of the existence of the oath, were asked to type a neutral sentence during the oath-taking stage.[20] Advisors were made aware that all advisors had to take the oath and were also informed of the asymmetry in information with respect to the investors. We compare behavior in the Oath-Private treatment with that in the No-Oath treatment to identify the importance of the self-image and audience-image costs of breaking an oath: since the instructions for investors were identical across these two treatments, the impact on the reputation of advisors to the investors as it relates to breaking the oath is the same (and nil) in both those treatments, so that differences in behavior across these treatments can be assigned to the joint private and audience costs of breaking an oath.

[18]As shown in Table A1 in Online Appendix A, the majority of previous experiments used voluntary oaths but in these studies, very few participants refused to sign it. We imposed that participants sign the oath to align ourselves with professional oaths that employees are typically required to sign when they are hired. In addition, it ensured that taking the oath was common knowledge. Participants were free to refuse participation in the experiment if they did not want to sign the oath; none did.

[19]The computer program checked that the words "I swear upon my honor", "honestly", and "truth" were correctly typed. We also checked the actual oaths typed by participants after the experiment. All were correctly typed, except for minor typos.

[20]This was done to avoid placing the investors in a situation in which they would wonder why other participants are typing while they were not.

Asymmetric Asynchronous ("Asym-Async") treatment: All participants in this treatment were assigned the role of investor. They were told that the instructions they would receive would be an *excerpt* of the instructions that were given to participants from a past session. Those instructions were those from the Oath-CK treatment and were identical to them, except for the removal of the section on the oath.[21] After reading those instructions, participants learned that in each period they were going to be randomly matched with advisors from those past sessions and would face the announcement of the advisor they would be matched with. In addition, their payment would be determined by Nature's draw for that particular advisor in that period.[22,23] To ensure that participants had enough opportunity to learn from experience, we extended the number of periods in this treatment to 27; as in the other treatments, this number was unknown to the participants. Comparing investment decisions in this treatment to those in the Oath-CK treatment allows us to study whether investors react mainly to the common-knowledge element of the oath or to their actual experience of advisors' honesty.

Symmetric Asynchronous ("Sym-Async") treatment: All participants in this treatment were assigned the role of investor. The procedure of the Sym-Async treatment was similar to that of the Asym-Async treatment, except that in part 2 the instructions explicitly integrated those used in the Oath-CK treatment including the part about the oath. Participants were then told that in each period they would be randomly matched with an advisor drawn from a *subset of advisors* from these past sessions. Unbeknownst to the participants, this subset of advisors was chosen such that the empirical distribution of announcements was as close as possible to that in the No-Oath treatment.[24] In other words, we created a situation where investors in Sym-Async treatment faced advisors who took a common-knowledge

[21] In the Oath-CK treatment, the reference to the honesty oath was placed at the very end of the instructions and we simply removed that section in the Asym-Async treatment. In our instructions we were very explicit that the instructions they were to receive were an "*excerpt*" of the prior ones. On the reasons for not considering withholding information about treatment manipulations as deception, see Hey (1998); Hertwig and Ortmann (2008); Charness et al. (2022).

[22] In each period and for each investor independently, the program randomly selected one of the 1386 announcements made by the 77 advisors who participated in the Oath-CK treatment and showed the investor that announcement. The card (star or blank) that determined the outcome of the investment in that period was the one actually drawn by Nature on that specific announcement in the past session. Thus, the Asym-Async treatment preserved the same empirical distribution of the advisors' announcements and Nature's draws than in the Oath-CK treatment. Using Kolmogorov-Smirnov test reveals no significant difference between the actual distribution of announcements in Asym-Async treatment (*ex-post*) and that in Oath-CK treatment ($p = 1.000$).

[23] Naturally, one other difference is that the investors' decisions in this treatment impact the payoff of no one else. We acknowledge that participants may take others' payoffs into account. However, Tergiman and Villeval (2023) showed that with the parameters used, social preferences do not play a large role in the Announcement Game. We provide further evidence on this in section 4.3.

[24] A Kolmogorov-Smirnov test reveals no significant difference between the distributions of announcements in Sym-Async treatment (*i.e.*, created by the subset of advisors in Oath-CK) and that in the No-Oath treatment ($p = 1.000$).

oath but behaved on average like those who did not take any oath. Comparing investment decisions in this treatment to those in the Oath-CK and No-Oath treatments allows us to examine whether the knowledge about the existence of the oath in itself is sufficient to persuade investors to invest and whether this changes over time as they experience untruthful announcements.

Belief elicitation: After completing the Announcement Game, all participants answered incentivized questions about their beliefs depending on their role. Investors had to estimate the proportion of honest announcements for each number of stars announced to them during the session (*i.e.,* their first order beliefs about the advisors' honesty). Advisors were asked to estimate both the investors' guesses about the proportion of honest announcements for each number of stars (*i.e.,* their second order belief) and the proportion of honest announcements for each number of stars made by all advisors in the session (*i.e.,* their first order belief). They earned a 0.50 Euro bonus if their answer in one (investors) or two (advisors) randomly selected guesses was accurate within an interval of 5% point. Finally, participants were told their earnings in each part and answered a questionnaire including socio-demographic information and questions on the motivation of choices in the experiment.

3.2 Procedures

The experiment was run at GATE-Lab, Lyon, France. We ran a total of 29 sessions: 7 sessions with 152 participants in the No-Oath treatment (76 participants in each role), 7 sessions with 154 participants in the Oath-CK treatment (77 in each role), 7 sessions with 146 participants in the Oath-Private treatment (73 in each role), 5 sessions with 77 participants in the Asym-Async treatment (all investors), and 3 sessions with 78 participants in the Sym-Async treatment (all investors).[25] All 607 participants were recruited via HRoot (Bock et al., 2014). Participants were primarily students from local engineering, business and medical schools.[26] The experiment was programmed using z-Tree (Fischbacher, 2007).

Upon arrival, participants randomly drew a ticket from an opaque bag which assigned them to terminals. Instructions were distributed at the beginning of the relevant part and read aloud. The average duration of a session was 80 minutes.

When designing the experiment, we avoided almost all interactions between the par-

[25] We based the number of observations on an ex ante power calculation. Assuming a Type-I error rate of 0.05 and a medium effect size (Cohen's d = 0.60), targeting 77 observations per role and per treatment would allow us to uncover the hypothesized effect of oath using a two-tailed Mann-Whitney ranksum test and achieve a power level of 0.95.

[26] See Tables C1 and C2 in Online Appendix C for summary statistics of the socio-demographic characteristics of advisors and investors.

ticipants and the experimenter, so as to minimize concerns that participants would have vis-à-vis how they are perceived by the experimenter. The instructions were audio-recorded; the experimenter left the room as soon as the questions about the instructions were answered in private (very few questions were asked and none were related to the oath); the computer program was launched and monitored by a technician. Importantly, participants were informed in advance that their earnings would not be paid in cash but wired directly on their bank account by bank transfer. Therefore, participants knew they would leave the room immediately at the end of a session without having to interact again with the experimenter. Because information for payment was entered by the participants on a separate file and transfered directly to the accountant of the institute without being communicated to the researchers, there was no way for the experimenters to tie decisions to a particular subject in an identifiable way.

Participants' earnings consisted of their payoffs from each part of the experiment: the payoff in the relevant scenario in the randomly chosen period (9 or 10) in the Truthful Announcement Game, the payoff from the randomly chosen round in the Announcement Game, and the payment from the belief elicitation. On average, participants earned 20 Euros (SD = 6.2), including a 7-Euro show-up fee.

4 Results

We begin by evaluating the overall impact of a common-knowledge honesty oath on advisor behavior, considering both detectable lies, which happen when an advisor sees fewer than three stars but announces three stars, and deniable lies, which happen when an advisor announces two stars but sees fewer than two stars, or announces one star while seeing none.[27] Next, we explore why oaths lead to changes in behavior and document the importance of the non-pecuniary costs of breaking an oath. We then turn to investors and explore whether change in investment behavior is due to the *knowledge* that an oath has been taken, or to the increase in honesty that investors actually experience. After analyzing the results from our laboratory treatments, we report on the data from our field survey in the Netherlands. We evaluate the extext to which the public is aware of the Dutch bankers' oath, which bankers are required to take, explore how this knowledge links to trust in the profession, and evaluate how the general population perceives oaths and their effectiveness.

[27]More precisely, regarding detectable lies, when fraudulently announcing three stars, the chances that the advisor's lie is detected ranges from 100% if the advisor actually saw zero stars, to 67% if the advisor saw one star among the three cards, to 33% if the advisor saw two stars among the three cards. With respect to deniable lies, Nature's draw will be consistent with the announcement, no matter whether the draw is of a blank card or of a card with a star, and thus the investor is unable to detect the lie.

In order to analyze the data, and unless otherwise specified, we use two-tailed rank-sum tests and tests of proportions as our main statistical tool to evaluate differences across treatments, using one observation per subject.[28] We further support our claims via the use of econometric analyses, using mainly random-effects linear probability models with clustering at the individual level, all placed in Online Appendix C.

4.1 Impact of the common-knowledge oath on lies

In order to evaluate the impact of the common-knowledge oath on lies we compare the No-Oath and the Oath-CK treatments.[29] Table 1 shows the frequency of detectable and deniable lies among advisors in the No-Oath and Oath-CK treatments. The fraction of advisors who always tell the truth increases significantly with the introduction of a common-knowledge oath, going from 9.2% in the No-Oath treatment to 74.0% in the Oath-CK treatment ($p < 0.001$). Looking at the population of advisors as a whole (top panel), it is clear that the reduction in lies after an oath is taken impacts both detectable and deniable lies: the relative frequency of the former decreases from 25.2% to 3.9% ($p < 0.001$) and the frequency of the latter decreases from 57.4% to 13.8% ($p < 0.001$).[30] The same patterns arise if we look at the fraction of advisors who engage in each type of lie at least once, as the "absolute frequency" data displayed in Table 1 shows.[31] In other words, the common-knowledge oath has a large and statistically significant impact on lying, for both detectable and deniable lies. The result regarding the reduction of deniable lies is particularly striking, as those lies are precisely those that are impossible for the investors to detect.

The oath also has an impact on those subjects who continue to lie despite it. Focusing on the subset of advisors who lie at least once when there is no oath (which represents 91.8% of advisors) and those who break the common-knowledge oath (26.0% of advisors), we note that the oath changes both the intensity and the nature of lies (lower panel of Table 1). While lies are still widespread among the advisors who do lie at least once, the

[28] We average choices within a subject and use this average choice as our unit of measure.

[29] Note here that lying costs (independently of those incurred by breaking an oath) exist in all treatments. We assume here that lying costs and the costs of breaking an oath enter the utility in an additive and separable manner. If one is to assume a more complicated view that these costs interact and that the partials are non-zero, then our design can tease out the joint impact of lying and oath-breaking costs.

[30] The Linear Probability Models reported in Tables C4 and C5 in Online Appendix C and estimating the probability of making a detectable and a deniable lie: a common-knowledge oath causes a reduction in detectable and deniable lies by about 20 and 40 percentage points, respectively, compared to when there is no oath ($p < 0.001$ for both; model (1)).

[31] In addition, we note no differences in time trends across the two treatments: when comparing the difference-in-differences between the relative frequencies of detectable and deniable lies between the first and second half of the experiment across the two treatments, the p-values are 0.819 and 0.850, respectively. This is confirmed by the regressions reported in Tables C4 and C5, and by visual inspection of the period-by-period frequencies over the course of the game, as displayed in Figures E1 and E2 in Online Appendix E.

Table 1: Detectable and deniable lies in the No-Oath and Oath-CK treatments

Treatments	No-Oath	Oath-CK	p-value
All advisors			
% Advisors who always tell the truth	9.2%	74.0%	$p < 0.001$
% Untrue announcements	52.3%	10.8%	$p < 0.001$
Relative frequency			
Detectable lies	25.2%	3.9%	$p < 0.001$
Deniable lies	57.4%	13.8%	$p < 0.001$
Absolute frequency			
Detectable lies	71.1%	11.7%	$p < 0.001$
Deniable lies	86.8%	23.4%	$p < 0.001$
Among advisors who lie at least once			
% Untrue announcements	57.6%	41.4%	$p = 0.010$
Relative frequency			
Detectable lies	27.8%	15.0%	$p = 0.007$
Deniable lies	63.2%	53.2%	$p = 0.302$
Absolute frequency			
Detectable lies	78.3%	45.0%	$p = 0.004$
Deniable lies	95.7%	90.0%	$p = 0.334$

Notes: Relative frequency statistics show the average frequency with which participants engage in any particular type of lie. Absolute frequency statistics show the proportion of advisors who make any particular type of lie at least once. The relative frequency of detectable lies corresponds to how often advisors who saw fewer than three stars announced three stars. The relative frequency of deniable lies corresponds to how frequently advisors who saw fewer than two stars reported two stars or announced one star while seeing none. This table presents data both for all advisors (top panel), and for the subset who make at least one lie (lower panel).

frequency of lies drops from 57.6% without the oath to 41.4% with the oath ($p = 0.010$). Interestingly, the difference in frequency only comes from detectable lies: we note a sharp reduction in the relative frequency of detectable lies (27.8% *vs.* 15%, $p = 0.007$), whereas deniable ones remain high and are no different across treatments for this subset of advisors (63.2% *vs.* 53.2%, $p = 0.302$). The same conclusion holds if we turn our attention to absolute frequencies (see bottom of Table 1).[32] This suggests that for the 26% of advisors who lie under oath, the impact of the oath is only partial and targeted to lies that would obviously make plain that the oath has been broken.

4.2 Non-pecuniary costs of breaking an oath

We now ask why the common-knowledge oath has such a large impact on advisor behavior. In order to answer this question, we turn to the non-pecuniary costs of breaking an oath. Advisors may incur three types of non-pecuniary costs when they break an oath. The first are costs that relate to how other players in the game perceive them. We call these *co-player-image costs*. The second are *self-image* costs, which come from how advisors may feel about themselves after breaking an oath. The third are *audience-image* costs, which come from concerns related to how an outside observer, such as the experimenter or God, may view them (in thinking about the audience, we share the view of Abeler et al. (2019) that "the audience differs from the other players in the game in that it takes no action, but rather serves as a player who may hold beliefs about any of the subjects after observing the subjects' reports").[33]

Co-player-image costs of breaking an oath

The impact of the co-player-image costs of breaking an oath can be extracted by comparing the Oath-CK and Oath-Private treatments. Indeed, the self-image costs of breaking an oath are held constant across these two treatments (as in both cases advisors swear an honesty oath). In addition, the comparison across these two treatments also controls for audience-image concerns, since audience observers are also aware of the oath and of the advisor's actions in both cases.

Impact of co-player-image costs: Table 2 summarizes advisor behavior across these two treatments. The top panel of Table 2 shows that when the oath is no longer common

[32] Models (3) in Tables C4 and C5 show that compared to the No-Oath treatment, detectable lies in the Oath-CK treatment are lower by 12 percentage points ($p = 0.091$), but only by 9 percentage points for deniable lies (and not significantly different).

[33] Alternatively, one may view self-image costs as costs that come from being one's own audience as in Dufwenberg and Dufwenberg (2018).

knowledge, the fraction of truth-tellers drops by 35 percentage points, going from 74.0% to 39.7% ($p < 0.001$). In other words, co-player-image costs account for about 35 of the 70 percentage point increase in honest advisors after the introduction of a common-knowledge oath relative to a situation in which no oath is present. We also note that while the common-knowledge oath sharply decreases the fraction of advisors who continue to lie compared with when the oath is privately taken, we find no statistical differences between the advisors who continue to lie in the Oath-CK treatment and those who continue to lie in the Oath-Private treatment, indicating that for these subjects any co-player-image costs induced by the oath are not sufficiently high to modify any behavior.

Table 2: Detectable and deniable lies in the Oath-CK and Oath-Private treatments

	Oath-CK	Oath-Private	p-value
All advisors			
% Advisors who always tell the truth	74.0%	39.7%	$p < 0.001$
% Untrue announcements	10.8%	27.7%	$p < 0.001$
Relative frequency			
Detectable lies	3.9%	9.6%	$p = 0.002$
Deniable lies	13.8%	37.9%	$p < 0.001$
Absolute frequency			
Detectable lies	11.7%	32.9%	$p = 0.002$
Deniable lies	23.4%	56.2%	$p < 0.001$
Among advisors who lie at least once			
% Untrue announcements	41.4%	46.0%	$p = 0.397$
Relative frequency			
Detectable lies	15.0%	15.9%	$p = 0.593$
Deniable lies	53.2%	62.8%	$p = 0.316$
Absolute frequency			
Detectable lies	45.0%	54.5%	$p = 0.479$
Deniable lies	90.0%	93.2%	$p = 0.660$

Notes: Relative frequency statistics show the average frequency with which participants engage in any particular type of lie. Absolute frequency statistics show the proportion of advisors who make any particular type of lie at least once. The relative frequency of detectable lies corresponds to how often advisors who saw fewer than three stars announced three stars. The relative frequency of deniable lies corresponds to how frequently advisors who saw fewer than two stars reported two stars or announced one star while seeing none. This table presents data both for all advisors (top panel) and for the subset who make at least one lie (lower panel).

Modeling co-player-image costs: These image costs are those that relate to the beliefs that the investors may hold regarding the advisor or their actions. One approach is to model co-player-image costs as a function of the investor's belief that the oath has been broken (the advisor feels judged on her action). A second approach is to think of co-player-image costs as coming from a desire to not be perceived as someone for whom it is easy to break an oath (the advisor feels judged as a person).[34]

If the co-player-image costs are specific to not wanting to be seen as having broken an oath, we should observe that the common-knowledge element of the oath would more sharply reduce "highly detectable" lies compared with "less detectable" ones, since the chances of being identified as a "confirmed oath-breaker" are higher when one announces three stars after observing none than after announcing three stars after observing two for example. We find that this is not the case,[35] despite the fact that the lower the true number of stars, the higher the benefit of making a detectable lie.[36] In fact, the absolute probability of making a detectable lie in the Oath-CK treatment is statistically independent of what the true number of stars is.[37] Therefore, the patterns in our data are incompatible with this modeling choice. Instead, though we cannot of course know how difficult a decision to lie is for a participant (morally speaking), our data are consistent with the idea that co-player-image costs stemming from the oath may come from not wanting to be seen as someone for whom it is easy to break an oath rather than simply someone who has or may have broken an oath.

Self-image and audience-image costs

The comparison between the No-Oath and Oath-Private treatments sheds light on the costs of breaking an oath that exist outside of the co-player-image costs, that is, those costs due to how individuals would feel or perceive themselves (self-image costs), or how an external observer would see them (audience-image costs), after breaking an oath, independently of how investors see them. Investors in the No-Oath and Oath-Private treatments face the same instructions since in the latter the oath remains private to the advisors. As a result, an advisor's image vis-à-vis an investor after a particular announcement and a particular payoff-determining draw cannot be impacted by beliefs that would be induced by knowledge

[34] The first approach is in line with Gneezy et al. (2018); Abeler et al. (2019); Khalmetski and Sliwka (2019). The second one is in line with Frankel and Kartik (2019).

[35] Relative to no oath, the reduction in extreme lies caused by the common-knowledge oath is not statistically different from those in high-risk lies ($p = 0.755$) and lower-risk lies ($p = 0.814$).

[36] We establish in Section 4.3 that investments in the Oath-CK treatment are monotonic in the number of stars announced.

[37] The probabilities of announcing three stars when the truth is zero, one, or two are 13.2%, 13.0% and 17.3%, respectively; these are not statistically different as the smallest pairwise p-value for signrank tests is 0.563).

of the oath. In other words, in the Oath-Private treatment there is no co-player-image cost of breaking the private oath vis-à-vis an investor, and any observed increase in honest behavior in the Oath-Private treatment relative to the No-Oath treatment is thus likely due to an increase in the self- or audience-image costs of breaking the oath.

Table 3: Detectable and deniable lies in the No-Oath and Oath-Private treatments

	No Oath	Oath-Private	p-value
All advisors			
% Advisors who always tell the truth	9.2%	39.7%	$p < 0.001$
% Untrue announcements	52.3%	27.7%	$p < 0.001$
Relative frequency			
Detectable lies	25.2%	9.6%	$p < 0.001$
Deniable lies	57.4%	37.9%	$p = 0.001$
Absolute frequency			
Detectable lies	71.1%	32.9%	$p < 0.001$
Deniable lies	86.8%	56.2%	$p < 0.001$
Among advisors who lie at least once			
% Untrue announcements	57.6%	46.0%	$p = 0.004$
Relative frequency			
Detectable lies	27.8%	15.9%	$p = 0.003$
Deniable lies	63.2%	62.8%	$p = 0.884$
Absolute frequency			
Detectable lies	78.2%	54.5%	$p = 0.008$
Deniable lies	95.6%	93.2%	$p = 0.568$

Notes: Relative frequency statistics show the average frequency with which participants engage in any particular type of lie. Absolute frequency statistics show the proportion of advisors who make any particular type of lie at least once. The relative frequency of detectable lies corresponds to how often advisors who saw fewer than three stars announced three stars. The relative frequency of deniable lies corresponds to how frequently advisors who saw fewer than two stars reported two stars or announced one star while seeing none. This table presents data both for all advisors (top panel) and for the subset who make at least one lie (lower panel).

Impact of self-image and audience-image costs. Before teasing apart the self- and audience-image costs, we first establish their joint impact. Table 3 summarizes advisor behavior across the No-Oath and Oath-Private treatments. In our sample, self- and audience-image costs of breaking the oath are high enough to fully impact 30% of participants, as the fraction of advisors who tell the truth goes from roughly 10% in the No-Oath treatment to about 40% in the Oath-Private treatment ($p < 0.001$). In other words, the costs of making an untruthful announcement are high enough to move 10% of subjects to tell the truth, while the self- and audience-image costs of breaking an oath are such that 30% additional subjects move to telling the truth. While a large majority of advisors continue to lie at least once when the oath is private, the private oath still impacts these advisors as they make significantly fewer detectable lies under private oath than when no oath is taken ($p = 0.003$ for relative frequency and $p = 0.008$ for absolute frequency).

Two sets of questions emerge. First, do these costs come from the advisor's beliefs as to what an outside audience may think, or are they more likely due to one's own conscience or character? Second, how should one think about modeling these costs?

Origins of costs: The post-experimental questionnaire helps inform us as to whether the non-co-player-image costs are primarily due to self-image or audience-image concerns. We focus on the stated motivations of the truth-tellers in the Oath-Private treatment.[38] None of the truth-tellers referenced religion or fear of God. None mentioned the experimenter explicitly (as it relates to how an experimenter may judge them). Four truth-tellers' stated motivation (about 17% of answers) did reference an external person: two described being truth-tellers because "I was asked to," where the reasonably natural completion of this passive voice would be "by the experimenter;" the other two described not wanting to corrupt the research question (assuming that the experimenter wanted them to tell the truth). The modal stated motivation (roughly 42% of answers) instead references placing weight on their honor or character as it relates to the oath, with narratives such as "I swore an oath not to [lie] and my honor is more important than money," "Breaking the oath wasn't worth the additional gain," "It is important to me to keep my word," and similar statements. Thus, our data point to self-image costs independent of any outside audience judgment as being the primary motive for truth-telling under private oath.[39]

[38] Twenty-nine advisors were truth-tellers, and all but four answered this part of the questionnaire.

[39] The very nature of experiments makes it impossible to design an experiment in which the oath would be fully unknowable to the experimenter, or to design an experiment in which one could control for judgment by God. Subjects may not actually sign the oath when it is no longer observable to the experimenter. Thus, an increase in lying could be an artefact of not signing an oath instead of a lower audience-image cost. In addition, designing an experiment where there is only some chance of an oath being requested would still lead to participants forming beliefs on how likely the experimenter knows an oath has been signed. In short, the hypothesis that participants care about an outside observer's beliefs about them in our setting is not falsifiable by its very nature. As such we are limited by the information that we collected via our questionnaire. We acknowledge that this is not as clean an identification that would rely on observed actions in a controlled experiment. We also acknowledge that it could be that some participants had several motivations guiding their choices and only chose to write the one that held the most importance to them.

Modeling self- or audience-image costs: We start by considering two models of costs: fixed costs (the cost incurred by breaking the oath does not depend on "how badly" the oath was broken), and costs that vary with the size of the lie that was made (for example, it is less costly to break an oath if one is to announce three stars when the truth is two than when the truth is one).

We first highlight that we cannot use data from consistent truth-tellers to better understand these costs, and so focus on subjects who have lied at least once.[40] Fixed costs are incompatible with the patterns we see in our data: indeed, if the costs of breaking an oath were fixed, then there should be no difference in terms of their impact on detectable and deniable lies and these should be impacted similarly. Instead we see a drop in the former and no difference in the latter: the private oath reduces detectable lies compared with the absence of an oath (27.8% versus 15.9%; $p = 0.003$) but has no impact on deniable lies (63.2% versus 62.8%; $p = 0.884$).

A second possibility is that instead of a fixed cost, the self- and audience-image costs of breaking the oath, c, depend solely on the "size" of the lie: $c = f(|m - \tau|)$ where f is increasing (the larger the difference between the truth and the announcement, the larger these costs of breaking the oath). We can largely rule this modeling choice out as well. Indeed, let us consider two subjects: one who is facing one star and considering announcing three stars, and the other who is facing zero stars and considering announcing two stars. Using our empirical values for the probabilities of investment at various announcement levels, the net benefits of breaking the oath relative to telling the truth are $150 - f(2)$ and $137 - f(2)$, respectively. Thus, we should find that the probability of announcing three stars after seeing one star is greater than the probability of announcing two stars after seeing none. However, we find the opposite, as the probability of the former is 14.5% and the latter 53.7% ($p < 0.001$, signrank test).

Thus, fixed costs or costs that depend only on the size of the lie cannot explain the patterns in our data. Our data suggest that an additional mechanism may be at play: subjects simply do not want to lie maximally if they have sworn to not lie at all: indeed, the probability of announcing three stars when seeing zero is not significantly different from the probability of announcing three stars when seeing one or when seeing two (15.5%, 14.5%, 18.5%, respectively; the smallest pairwise p-value for signrank tests is 0.172), despite the fact that lying maximally would yield a higher return than the truth when the truth is zero stars than when it is one or two (see Section 4.3).[41] In addition, and in support of the argument

[40] In addition, we note that almost all such subjects in the No-Oath treatment who make detectable lies also make deniable ones. This rules out that subjects who lie in the Oath-Private treatment would only be subjects who exclusively make detectable lies.

[41] Recall that in the Oath-Private treatment, investors are unaware of the oath. Thus, the self-image costs of

that "lying maximally" costs are due to breaking the oath, we point out that in the No-Oath treatment, maximal lies are twice as common, and the probabilities of lying maximally in the presence of an oath (whether private or common-knowledge) are significantly lower than those in the No-Oath treatment (the largest p-value is 0.038), showing that the oath itself changes behavior.

In summary, the mechanism through which an oath can be effective is three-fold and only an oath taken with common knowledge can raise all of these costs. In our data a common-knowledge oath leads close to 75% of advisors to tell the truth, compared with only 40% when the oath is taken privately, and 10% when there is no oath.

4.3 Oaths and investors

Impact on investment and earnings

We now turn to investor behavior and look at the impact of the common-knowledge oath, both on overall investment and on investment after two- or three-star announcements. Table 4 shows the average investment rates as well as how investors react to various levels of announcements in both the Announcement Game and the Truthful Announcement Game. Comparing behavior across these two games and across treatments allows us to understand how credible various announcements are and how the common-knowledge oath impacts credibility and investment behavior, accounting for the investors' baseline investment rates in the absence of asymmetric information.[42]

There are significantly more investments after two- and three-star announcements under the common-knowledge oath than without the oath.[43] The investment rate after a three-star announcement increases by almost 20 percentage points after the introduction of a common-knowledge oath, going from 77.8% to 97.7% ($p < 0.001$). The corresponding statistics for investments after a two-star announcement are 70.5% and 85.8% ($p < 0.001$). The data also show that investors find announcements more credible in the Oath-CK treatment relative to the No-Oath treatment. Indeed, the difference in investment rates between the Truthful Announcement and Announcement Games is significantly smaller in the Oath-CK than in the No-Oath treatment for two and three stars ($p < 0.001$ in both cases), showing that the

breaking an oath and lying maximally are not related to the co-player-image costs of being identified as someone who lies despite signing an oath.

[42] As a complement, Online Appendix D provides additional information on the investors' end-of-experiment first-order beliefs about the credibility of the announcements.

[43] The Linear Probability model of the likelihood of investing reported in Online Appendix Table C6 indicates an increase in investment by 17 percentage points in the Oath-CK treatment compared to No-Oath ($p < 0.001$; model (1)).

Table 4: Investment rates in the Announcement and Truthful Announcement Games

		No-Oath	Oath-CK
0 stars announced	% Announcement Game	9.5%[1] (21)	1.7% (59)
	% Truthful Ann. Game	0.0%	2.6%
1 star announced	% Announcement Game	14.5% (54)	16.8% (77)
	% Truthful Ann. Game	19.7%	16.9%
2 stars announced	% Announcement Game	70.5% (76)	85.8% (77)
	% Truthful Ann. Game	98.7%	98.7%
3 stars announced	% Announcement Game	77.8% (76)	97.7% (72)
	% Truthful Ann. Game	100.0%	98.7%
Average investment rate		65.8%	56.9%
Average investor earnings		149.7	159.9

Notes: The table displays the fraction of times participants invested in the Announcement and Truthful Announcement Games. Numbers of participants faced with a particular announcement at least once are in parentheses. [1] While this percentage may seem relatively high, this represents a small number of participants, as only 21 investors saw a zero-star announcement in this treatment, and of these two invested.

gap closes when the common-knowledge oath is introduced and, in parallel, that investors are more suspicious about their advisors' honesty in the baseline than in the Oath-CK treatment. As investment rates rise sharply with the introduction of the common-knowledge oath, investors move closer to the choices they would have made if information were symmetric. In fact, when three stars are announced the gap across the Truthful Announcement Game and the Announcement Game narrows to one percentage point (97.7% in the Announcement Game versus 98.7% in the Truthful Announcement Game).[44]

Finally, since there are fewer two- and three-star announcements under oath than in the baseline, overall average investment levels are slightly lower under oath, but the proportion of "good investments" compensates for this drop and leads to higher earnings for investors. Indeed, as announcements are more frequently truthful and credible, investors avoid investing when the truth is zero stars, and, at the same time, are more likely to invest when the actual number of stars is two or three.[45] This leads to an increase in investor earnings from 149.7

[44]Comparing these investment rates with investment rates under the Truthful Announcement Game in the absence of any information (period 10 in part 1), we see that in fact under the No-Oath treatment, two-star announcements are treated as uninformative, while three-star announcements do carry some information ($p = 0.550$ and $p < 0.001$, respectively – signrank tests), while both types of announcements under Oath-CK are treated as containing information ($p = 0.005$ and $p < 0.001$, respectively – signrank tests).

[45]The proportions of "good investments" (*i.e.,* made when the announcement is true) on two- and three-star

tokens without the oath to 159.9 tokens under oath ($p = 0.016$).[46]

Why do investors invest more under oath?

We now ask what role knowledge of the oath plays in investor behavior: are investors reacting to their experience with more truthful advisors, or are they reacting to the knowledge of the oath *per se*? In other words, we ask whether knowledge of the oath is *necessary* for investors to trust the announcements, and whether that knowledge is *sufficient* for investors to think the announcements are credible?

A necessary condition? In order to answer the first question, we turn to the Asym-Async treatment. In this treatment, all participants have the role of investors and they are faced with the announcements from the advisors in the Oath-CK treatment (who behaved predominantly honestly), without being aware that those advisors had sworn an oath. Columns (1) and (2) in Table 5 place side-by-side the behavior of investors in the Asym-Async and Oath-CK treatments, focusing on the last nine periods of play to allow participants the opportunity to learn, particularly in the Asym-Async treatment.[47]

Table 5 shows that average investment rates are no different across these two treatments (56.1% in Asym-Async *vs.* 56.0% in Oath-CK, $p = 0.860$) (see also the non-significant Wald test at the bottom of Table C6 in Online Appendix C).[48] The gap in investment rates after zero-, one-, two- or three-star announcements across treatments is always small in magnitude and statistically not significant: investment choices in the Asym-Async treatment are no different than those from investors who were aware of the oath. In addition, while the investment rate decreases over time when there is no oath, such a time trend is not observed in the Oath-CK or Asym-Async treatments.[49]

announcements in the Oath-CK treatment are 83.3% and 83.8% respectively. The corresponding figures in the No-Oath treatment are 50.2% and 33.2%, which are significantly lower ($p < 0.001$ for both).

[46] For advisors the lower investment rate under oath naturally leads to a significant drop in their average earnings from 161.6 tokens in No-Oath to 143.9 tokens in Oath-CK ($p < 0.001$), and 145.1 tokens in Oath-Private ($p = 0.005$), with no significant difference between the two oath treatments ($p = 0.608$).

[47] Recall that in the Oath-CK treatment, there are a total of 18 periods so that we focus on the last half of the game, and in the Asym-Async treatment a total of 27 so that we focus on the last third. Our conclusions are robust to focusing on the second half of all periods in Asym-Async, and are largely unchanged if we instead restrict the analysis to the first 18 periods in order to match the number of periods between the Oath-CK and Asym-Async treatments (the only difference is that, while close in magnitude, there is a statistically significant difference in investment rates across the two sets of data when three stars are announced: 92.1% *vs.* 97.7%, $p = 0.076$). Conclusions are similar if we instead look at all the periods in each treatment without any restriction.

[48] The fact that investment decisions are similar across the Oath-CK and Asym-Async treatments, even though they only impact investor payoffs in the latter, indicates that in our setting social preferences may be outweighed by other considerations.

[49] This can be seen from comparisons between investment rates in the first nine periods and the last nine (57.9% *vs.* 56% in Oath-CK; 55.4% *vs.* 56.1% in Asym-Async). The difference is insignificant in the Oath-CK treatment ($p = 0.409$, signrank test), as well as in the Asym-Async treatment ($p = 0.977$, signrank test), whereas there is a

Table 5: Investment rates in the Asym-Async, Oath-CK, and Oath-Private treatments

	Asym-Async (1)	Oath-CK (2)	Oath-Private (3)	*p-value* *(1) vs. (2)*	*(1) vs. (3)*
0 stars announced	1.1%	2.3%	1.6%	$p = 0.966$	$p = 1.000$
1 star announced	11.3%	15.9%	10.5%	$p = 0.455$	$p = 0.487$
2 stars announced	84.7%	82.8%	70.4%	$p = 0.388$	$p < 0.001$
3 stars announced	95.2%	97.2%	88.8%	$p = 0.587$	$p = 0.068$
Average investment rate	56.1%	56.0%	56.9%	$p = 0.860$	$p = 0.513$

Notes: The table shows the fraction of times participants in the role of investors invested in the Announcement Game, looking at the last nine periods in order to allow participants to gain experience.

These data have two implications. First, the increase in investment after two- and three-star announcements in the Async-Asym treatment compared to the No-Oath treatment is robust to facing a small fraction of lies. Recall that just over a quarter of advisors lie in the Oath-CK treatment despite having sworn not to. In other words, although investors in the Asym-Async treatment do face some dishonest behavior, they still invest at very high rates. Second, making the oath common knowledge is not a necessary condition for its impact to be effective on investors: trusting announcements can be learnt even in a setting with asymmetric information and the possibility of deniable lies. This is further supported by the data from the Oath-Private treatment where investors are unaware of the oath (as is the case in the Asym-Async treatment) but experience different levels of truthfulness (as advisors in Oath-Private are significantly less honest than those that investors face in the Asym-Async treatment). The comparison of columns (1) and (3) in Table 5 shows that investment rates in the last nine periods in the Asym-Async treatment are higher than in the Oath-Private treatment after two- and three-star announcements ($p < 0.001$ and $p = 0.068$, respectively).[50] In other words, knowledge about the oath in and of itself is not necessary for trust to develop.

declining trend in the No-Oath treatment (69.2% *vs.* 62.4%, $p = 0.009$, signrank test). Table C7 in Online Appendix C reports random-effects regressions of the probability of investing on a time trend, with clustering at the individual level. The p-value of the period variable is equal to 0.843 in Oath-CK and 0.532 in Asym-Async, whereas it is equal to 0.022 in No-Oath.

[50] Table C6 in Online Appendix C reports the estimates of random-effect Linear Probability models of the probability of investing, considering all periods and with the No-Oath treatment as the reference category. Whereas the coefficient of the Oath-Private treatment dummy is not significant, the coefficient of the Asym-Async treatment dummy is significant (0.166, $p < 0.001$) and does not differ ($p = 0.894$) from the coefficient of the Oath-CK treatment dummy (0.171, $p < 0.001$).

A sufficient condition? If knowledge of the oath is not necessary, is it sufficient to convince investors to trust the announcements? To answer this question, we use data from the Sym-Async treatment. In this treatment, all participants had the role of investors, were given the instructions of the Oath-CK treatment (including the oath), and were told that they would face announcements from a *subset* of advisors in the Oath-CK treatment. The subset of advisors was selected so that the distribution of announcements would closely resemble the distribution of announcements in the No-Oath treatment. Table 6 shows investment behavior in the last nine periods of play in the Sym-Async, Oath-CK, and No-Oath treatments.

Table 6: Investment rates in the Sym-Async, Oath-CK and No-Oath treatments

	Sym-Async (1)	Oath-CK (2)	No-Oath (3)	*p-value* *(1) vs. (2)*	*(1) vs. (3)*
0 stars announced	0.0%	2.3%	0.0%	$p = 1.000$	$p = 1.000$
1 star announced	6.6%	15.9%	14.8%	$p = 0.064$	$p = 0.302$
2 stars announced	64.1%	82.8%	65.9%	$p = 0.002$	$p = 0.947$
3 stars announced	80.2%	97.2%	76.1%	$p < 0.001$	$p = 0.513$
Average investment rate	65.1%	56.0%	62.4%	$p = 0.010$	$p = 0.430$

Notes: The table shows the fraction of times participants in the role of investors invested in the Announcement Game, looking at the last nine periods in order to allow participants to gain experience.

Investor behavior in the Sym-Async treatment differs significantly from that in the Oath-CK treatment, but matches that of the No-Oath treatment. This is the case if we look at the overall average investment rates, but also if we look at investment rates after particular announcements (see the last two columns that show the p-values comparing treatments). These patterns show that knowledge of the oath is not sufficient to lead investors to trust advisors. Instead, investors are influenced by their own experience: when they know about the existence of an oath but realize that advisors have perjured themselves, they become less likely to invest: although investors in the Sym-Async treatment receive the same instructions as investors in the Oath-CK treatment, they behave as investors in the No-Oath treatment.[51]

Does knowledge that an oath has been taken then have no impact at all? What we have shown above is that if it does have an impact, it certainly is not persistent over time. However, it could be that knowledge of the oath *per se* changes initial attitudes and be-

[51] The investment rate also tends to decrease over time in the Sym-Async treatment (from 77.2% in the first nine periods to 65.1% in the last nine periods, $p < 0.001$), as in the No-Oath treatment, and in contrast with the Oath-CK treatment where it is stable, as previously shown. In the random-effects Linear Probability models of the probability of investing reported in Table C7 in Online Appendix, the p-value of the Period variable is <0.001 in Sym-Async, whereas it is equal to 0.022 in No-Oath and 0.843 in Oath-CK).

liefs. While we are very cautious in making claims regarding what may be going through our subjects' minds, our conjecture, and what our data suggest, is that initial behavior is not necessarily influenced by the knowledge that an oath has been taken. We support this claim by comparing first-period behavior, that is before any experience of play, across the Sym-Asym and Asym-Async treatments, which differ only in whether investors knew of the oath. We find no difference in how they react to announcements (the lowest p-value in all pairwise comparisons in a test of proportions for each type of announcement is 0.106). Had knowledge of the oath been a sufficient condition to believe it would lead to higher levels of truthfulness, then we would have seen higher investment rates for both two- and three-star announcements in the Sym-Async treatment.[52]

In summary, from the investor's stand-point, an oath is neither a sufficient nor a necessary condition for trust to be established, and the impact that we observe in the Oath-CK treatment stems from the difference in advisor behavior that it results in, which itself changes investor behavior.

4.4 Field survey

In this section, we report the findings of an online survey conducted via Prolific with 198 respondents in the Netherlands. Through this survey, we aim to evaluate the extent of awareness of the Dutch bankers' oath, and explore how this knowledge links to the trust people have in the banking profession.[53]

The results of the survey complement our laboratory findings. While 69% of the respondants indicate that "in general people can be trusted," this trust does not extend to bankers for who the average score on a 1-5 Likert scale was 2.55, the lowest of all the professions that we inquired about.[54] This may be surprising for a country in which a banker oath is compulsory. In fact, the survey also points to the fact that while the oath is compulsory, the awareness of it is limited: only 53% of our respondents are either sure or think it is likely that such an oath exists. The awareness is equally limited if one looks at those with a financial advisor (59%).[55] The lack of awareness may hinder the public's trust in the banking sector

[52] This is further supported by similar comparisons across the Oath-Private and Oath-CK treatments, as well as between the Oath-CK and Sym-Async treatments. In particular, there is no difference in how investors react to two-star announcements in the first period, showing that experience is necessary for trust to be established.

[53] The full questionnaire and a more detailed discussion of the responses can be found in Online Appendix F.

[54] The average score for other professions was 4.62 for firefighters, 3.96 for doctors, 3.87 for judges, and 2.61 for members of the parliament.

[55] To better appreciate this fraction we conducted a similar survey in France. Bankers in France do not swear such oaths yet 28% of the respondents were either sure or thought it was likely that such an oath existed in France. This indicates that the awareness statistics in the Dutch survey are likely an over-estimate.

as there is a positive correlation between being aware of the oath and the level of trust in bankers.[56] Indeed, respondents who are aware of the oath hold a higher level of trust in bankers (Spearman's correlation coefficient = 0.15, $p = 0.036$). Our experiment shows that for investors, being aware of the oath is not a necessary condition for trust since participants learn from their experience in the market. This element is also reflected in our survey, as about 60% of the respondents with high trust in bankers indicate that it is because of the good experiences they have had with their bankers and banks. Nevertheless, the survey suggests that better awareness of the oath would improve trust in this industry.

Finally, our survey reveals that more than 80% of our respondents support making bankers oath mandatory across Europe and 60% believe that it would help make clients' interests more salient, alleviating one of the origins of mistrust. This is of course not a "cure-all", and we note that this would certainly require interventions on the financial advisors' side as well, such as reminders of the existence of the oath (our survey is silent on this point but Weitzel and Kirchler (2023) have shown in their audit study that with such moral nudges financial advisors were less likely to priorize the bank's interests).

5 Discussion and conclusion

As our survey and many other works have shown, the financial industry is regularly pointed out as a sector with low trust between companies and customers. In reaction to such crisis of trust, a few countries (the Netherlands and Belgium) have introduced a legally compulsory banker oath of good conduct that all employees have to sign. While its effectiveness would be hard to assess with standard data from the field, our experimental study shows that honesty oaths can have a large impact on truthful announcements in settings with asymmetric information. It supports the view that one can take advantage of preferences for promise-keeping to steer individuals towards more honest behavior, particularly when preferences for truth-telling may be outweighed by financial gain. Oaths discourage detectable lies and, more surprisingly, deniable lies. Our results are all the more striking as they have been obtained in a setting where there is no direct sanction for breaking the oath.

We establish the importance of three different image costs of breaking an oath in explaining how an oath influences behavior. At least half of the participants in the role of a financial advisor who react to the common-knowledge oath are impacted through image

[56] This lack of awareness was already highlighted as problematic by an independent commission evaluating in 2012 the original oath introduced in 2010. The Chairman of the banking association reported that while banks had abided by the code few people in the public were aware of it. See https://www.cnbc.com/2014/02/11/should-bankers-swear-an-oath-to-god.html, accessed on August 8, 2023. Our short survey indicates that this problem has not been solved and that a better communication campaign might be needed.

costs as they relate to the investors, even in a context in which anonymity is guaranteed and relationships are short-termed. For the remainder, the self-image and audience-image costs of breaking the oath are sufficient to move them to truthful behavior, which is consistent with the notion that both breaking an oath and lying oppose Kantian categorical imperatives. Removing universal awareness of the oath increases fraud immediately by reducing the advisors' reputational costs toward the investors. It is less likely that the reputational costs toward the investors originate from the preference for not being identified as a liar (as even deniable lies decrease under a common-knowledge oath) but from the preference for not being seen as someone for whom it is likely easy to lie, that is, someone with a low private cost of lying (in the spirit of Frankel and Kartik, 2019).

Our study also reveals that the more honest behavior of financial advisors that results from the oath leads to better investor outcomes, as investment profiles move closer to what investors would do under symmetric information. We show that the oath is neither sufficient nor necessary for investors to trust advisors: investors mainly rely on their past experience with advisors when making investment decisions. This element is also reflected in our Dutch survey as about 60% of respondents with high trust in bankers indicated that it was because of the good experiences they have had with their bankers and banks.

Overall, our results support that oaths might be an inexpensive and effective policy tool that would help move towards solving the crisis of confidence in the banking industry. This would be all the more important as a majority of households requests the advice of financial advisors for investing, and previous literature has shown that the delegation of financial decisions to advisors is influenced by both the investors' level of trust in the profession (Gennaioli et al., 2015; Germann et al., 2018) and a lack of financial literacy (Calcagno and Monticone, 2015; Stolper, 2018), in a context where financial products are more and more complex (Celerier and Vallee, 2017). Introducing a bankers' oath might be beneficial for investors and the profession more broadly, but this should be accompanied by regular information campaigns with both bank employees and the general public to ensure the universal awareness of the oath.

Oaths are obviously not the only approach that may work. A change in the incentive structure may be needed to have a larger impact on behavior through increased responsibility. Another problematic element in this industry is, as Egan et al. (2019) have shown, that even when criminal behavior is discovered and advisors are fired from their firm, they are often re-hired by other firms. In other words, the market undoes the firm-level punishment. An advantage of introducing an oath at the industry level is that it creates a sort of reputational public good at that level. In addition of better auditing, being able to track and identify past offenders across firms would increase the costs of committing fraud, by making future employment more difficult and playing on one's reputational costs as committing fraud would come more explicitly at the cost of an industry-wide "bad" reputation.

Avenues for future work are numerous. Developing more precise models and better understanding the structure of the image costs of breaking an oath would be valuable across a number of settings (the financial setting and more generally credence good markets among others, but also individual decision-making settings). Better understanding the empirical size of these costs and how they evolve over time may also be particularly useful to predict in what settings these costs outweigh the benefits of breaking an oath. And of course being able to predict who may or may not be impacted by an oath would also be helpful.

References

Abeler, J., D. Nosenzo, and C. Raymond (2019). Preferences for Truth-Telling. *Econometrica 87*(4), 1115–1153.

Akin, Z. (2022). Playing the victim behavior: An experimental study. Technical report, American University in Dubai.

Anagol, S., S. Cole, and S. Sarkar (2017). Understanding the advice of commissions-motivated agents: evidence from the Indian life insurance market. *Review of Economics and Statistics 99*(1), 1–15.

Andrews, C. R. (2009). The Lawyer's Oath: Both Ancient and Modern. *Georgetown Journal of Legal Ethics 22*, 3.

Babin, J. J. and H. S. Chauhan (2023). Show no quarter: combating plausible lies with ex-ante honesty oaths. *Journal of the Economic Science Association 9*, 66–76.

Babin, J. J., H. S. Chauhan, and F. Liu (2022). You Can't Hide Your Lying Eyes: Honesty Oaths and Misrepresentation. *Journal of Behavioral and Experimental Economics 98*, 101880.

Babri, M., B. Davidson, and S. Helin (2021). An updated inquiry into the study of corporate codes of ethics: 2005–2016. *Journal of Business Ethics 168*, 71–108.

Beck, T. (2021). How the honesty oath works: Quick, intuitive truth telling under oath. *Journal of Behavioral and Experimental Economics 94*, 101728.

Beck, T., C. Buhren, B. Frank, and E. Khachatryan (2020). Can honesty oaths, peer interaction, or monitoring mitigate lying? *Journal of Business Ethics 163*(3), 467–484.

Boatright, J. R. (2013). Swearing to be Virtuous: The Prospects of a Banker's Oath. *Review of Social Economy 71*(2), 140–165.

Bock, O., I. Baetge, and A. Nicklisch (2014). hroot: Hamburg Registration and Organization Online Tool. *European Economic Review 71*, 117–120.

Cagala, T., U. Glogowsky, and J. Rincke (2021). Detecting and Preventing Cheating in Exams: Evidence from a Field Experiment. *Journal of Human Resources 58*(3).

Cagala, T., U. Glogowsky, J. Rincke, and S. Schudy (2023). Commitment Requests Do Not Affect Truth-Telling in Laboratory and Online Experiments.

Calcagno, R. and C. Monticone (2015). Financial literacy and the demand for financial advice. *Journal of Banking and Finance 50*, 363–380.

Carlsson, F., M. Kataria, A. Krupnick, E. Lampi, Löfgren, P. Qin, T. Sterner, and S. Chung (2013). The truth, the whole truth, and nothing but the truth–a multiple country test of an oath script. *Journal of Economic Behavior & Organization 89*, 105–121.

Casella, A., N. Kartik, L. Sanchez, and S. Turban (2018). Communication in context: Interpreting promises in an experiment on competition and trust. *Proceedings of the National Academy of Sciences of the USA 115*(5), 933–938.

Celerier, C. and B. Vallee (2017). Catering to investors through security design: Headline rate and complexity. *The Quarterly Journal of Economics 132*(3), 1469–1508.

Charness, G. and M. Dufwenberg (2006). Promises and partnership. *Econometrica 74*(6), 1579–1601.

Charness, G., A. Samek, and J. van de Ven (2022). What is considered deception in experimental economics? *Experimental Economics 25*, 385–412.

Davis, B. J. and T. Jaber-Lopez (2022). Do voluntary commitment mechanisms improve welfare? The effect of mandatory and voluntary oaths in a social dilemma. *Bulletin of Economic Research 75*(2), 525–540.

de Bruin, B. (2016). Pledging integrity: Oaths as forms of business ethics management. *Journal of Business Ethics 136*(1), 23–42.

Di Bartolomeo, G., M. Dufwenberg, S. Papa, and F. Passarelli (2019). Promises, expectations and causation. *Games and Economic Behavior 113*, 137—-146.

Douglass, J. E. (1995). Between Pettifoggers and Professionals: Pleaders and Practitioners and the Beginnings of the Legal Profession in Colonial Maryland 1634-1731. *American Journal of Legal History 39*(3), 359–384.

Dufwenberg, M. and M. A. Dufwenberg (2018). Lies in disguise – a theoretical analysis of cheating. *Journal of Economic Theory 175*, 248 – 264.

Ederer, F. and F. Schneider (2022). Trust and promise over time. *American Economic Journal: Microeconomics 14*(3), 304–320.

Ederer, F. and A. Stremitzer (2017). Promises and expectations. *Games and Economic Behavior 106*, 161—-178.

Efing, M., H. Hau, P. Kampkötter, and J. Steinbrecher (2015). Incentive pay and bank risk-taking: Evidence from austrian, german, and swiss banks. *Journal of International Economics 96*(51), S123–S140.

Egan, M., G. Matvos, and A. Seru (2019). The Market for Financial Adviser Misconduct. *Journal of Political Economy 127*(1), 233–295.

Ellingsen, T. and M. Johannesson (2004). Promises, threats and fairness. *The Economic Journal 114*(495), 397–420.

Fecht, F., A. Hackethal, and Y. Karabulut (2018). Is proprietary trading detrimental to retail investors? *The Journal of Finance 73*(3), 1323–1361.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Fischbacher, U. and F. Föllmi-Heusi (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association 11*(3), 525–547.

Foerster, S., J. T. Linnainmaa, B. T. Melzer, and A. Previtero (2017). Retail Financial Advice: Does One Size Fit All? *The Journal of Finance 72*(4), 1441–1482.

Frankel, A. and N. Kartik (2019). Muddled information. *Journal of Political Economy 127*(4), 1739–1776.

Gennaioli, N., A. Shleifer, and R. Vishny (2015). Money doctors. *The Journal of Finance 70*(1), 91–114.

Germann, M., B. Loos, and M. Weber (2018). Trust and delegated investing: A money doctors experiment. *CEPR Discussion Paper 12984*.

Gneezy, U., A. Kajackaite, and J. Sobel (2018). Lying Aversion and the Size of the Lie. *American Economic Review 108*(2), 419–453.

Gravelle, H. (1993). Product price and advice quality: implications of the commission system in life assurance. *Geneva Papers on Risk and Insurance Theory 18*, 31–53.

Guiso, L., P. Sapienza, and L. Zingales (2008). Trusting the Stock Market. *The Journal of Finance 63*(6), 2557–2600.

Heese, J., G. Pérez-Cavazos, and C. D. Peter (2023). When Executives Pledge Integrity: The Effect of the Accountant's Oath on Firms' Financial Reporting. *The Accounting Review*, 1–28.

Heinicke, F., S. Rosenkranz, and U. Weitzel (2019). The effect of pledges on the distribution of lying behavior: An online experiment. *Journal of Economic Psychology 73*, 136–151.

Hergueux, J., N. Jacquemet, S. Luchini, and J. F. Shogren (2022). Leveraging the Honor Code: Public Goods Contributions under Oath. *Environmental and Resource Economics 81*(3), 591–616.

Hertwig, R. and A. Ortmann (2008). Deception in experiments: Revisiting the arguments in its defense. *Ethics & Behavior 18*(1), 59–92.

Hey, J. (1998). Experimental economics and deception: A comment. *Journal of Economic Psychology 19*(3), 397–401.

Inderst, R. and M. Ottaviani (2012a). Competition through commissions and kickbacks. *American Economic Review 102*(2), 780–809.

Inderst, R. and M. Ottaviani (2012b). How (not) to pay for advice: A framework for consumer financial protection. *Journal of Financial Economics 105*(2), 393–411.

Ismayilov, H. and J. Potters (2016). Why do promises affect trustworthiness, or do they? *Experimental Economics 19*(2), 382—-393.

Jacquemet, N., A. James, S. Luchini, J. Murphy, and J. Shogren (2021). Do truth-telling oaths improve honesty in crowd-working? *PLoS ONE 16*.

Jacquemet, N., A. James, S. Luchini, and J. F. Shogren (2017). Referenda Under Oath. *Environmental and Resource Economics 67*(3), 479–504.

Jacquemet, N., R.-V. Joule, S. Luchini, and J. F. Shogren (2013). Preference elicitation under oath. *Journal of Environmental Economics and Management 65*, 110–132.

Jacquemet, N., S. Luchini, A. Malézieux, and J. F. Shogren (2020). Who'll stop lying under oath? Empirical evidence from tax evasion games. *European Economic Review 124*, 103369.

Jacquemet, N., S. Luchini, J. Rosaz, and J. F. Shogren (2019). Truth Telling Under Oath. *Management Science 65*(1), 426–438.

Jacquemet, N., S. Luchini, J. Rosaz, and J. F. Shogren (2021). Can we commit future managers to honesty? *Frontiers in Psychology*, 2785.

Jacquemet, N., S. Luchini, J. F. Shogren, , and A. Zylbersztejn (2023). Commitment to the truth creates trust in market exchange: Experimental evidence. Technical report, GATE-LSE working paper series.

Jacquemet, N., S. Luchini, J. F. Shogren, and A. Zylbersztejn (2018). Coordination with communication under oath. *Experimental Economics 21*(3), 627–649.

Johnson, E., S. Meier, and O. Toubia (2016). Leaving money on the kitchen table: Exploring sluggish mortgage refinancing using administrative data, surveys, and field experiments. *NA - Advances in Consumer Research 44*, 179–184.

Kemper, N., J. Popp, and R. Nayga (2020). A query theory account of a discrete choice experiment under oath. *European Review of Agricultural Economics 47*(3), 1133–1172.

Khalmetski, K. and D. Sliwka (2019). Disguising lies—image concerns and partial lying in cheating games. *American Economic Journal: Microeconomics 11*(4), 79–110.

Koessler, A.-K. (2022). Pledges and how social influence shapes their effectiveness. *Journal of Behavioral and Experimental Economics 98*, 101848.

Koessler, A.-K., L. Page, and U. Dulleck (2021). Public cooperation statements. *Journal of Economic Interaction and Coordination 16*(4), 747–767.

Koessler, A.-K., B. Torgler, L. P. Feld, and B. S. Frey (2019). Commitment to pay taxes: Results from field and laboratory experiments. *European Economic Review 115*, 78–98.

Kowaleski, Z. T., A. G. Sutherland, and F. W. Vetter (2020). Can ethics be taught? Evidence from securities exams and investment adviser misconduct. *Journal of Financial Economics 138*(1), 159–175.

Krüger, M. P. (2016). Can an oath reduce dishonesty and overconfidence for bankers? Technical report, Mimeo.

Lang, M. and S. Schudy (2020). (dis)honesty and the value of transparency for campaign promises. Technical report, CESifo Working Paper.

Loonen, T. and M. R. Rutgers (2017). Swearing to be a good banker: Perceptions of the obligatory banker's oath in the Netherlands. *Journal of Banking Regulation 18*(1), 28–47.

Mullainathan, S., M. Noeth, and A. Schoar (2012). The market for financial advice: An audit study. *NBER Working Paper Series 17929*.

Naka, M. and H. Naoi (1995). The effect of repeated writing on memory. *Memory & Cognition 23*(2), 201–212.

Peer, E. and Y. Feldman (2021). Honesty pledges for the behaviorally-based regulation of dishonesty. *Journal of European Public Policy*, 1–21.

Pool, V. K., C. Sialm, and I. Stefanescu (2016). It pays to set the menu: Mutual fund investment options in 401 (k) plans. *Journal of Finance 71*(4), 1779–1812.

Prima, R. A., S. Feeny, R. Hoffmann, and E. Satriawan (2020). Threats, Pledges, and Asset Misreporting: A Framed Field Experiment in Indonesia. *TNP2K Working Paper 57-e/2020*.

Sapienza, P. and L. Zingales (2012). A Trust Crisis. *International Review of Finance 12*(2), 123–131.

Schild, C., D. W. Heck, K. A. Ścigała, and I. Zettler (2019). Revisiting REVISE: (Re)Testing unique and combined effects of REminding, VIsibility, and SElf-engagement manipulations on cheating behavior. *Journal of Economic Psychology 75*, 102161.

Serra Garcia, M., E. E. C. van Damme, and J. J. M. Potters (2013). Lying about what you know or about what you do? *Journal of the European Economic Association 11*(5), 1204–1229.

Skinner, C. H., T. McLaughlin, and P. Logan (1997). Cover, copy, and compare: A self-managed academic intervention effective across skills, students, and settings. *Journal of Behavioral Education 7*(3), 295–306.

Stolper, O. (2018). It takes two to tango: Households' response to financial advice and the role of financial literacy. *Journal of Banking and Finance 92*, 295–310.

Tergiman, C. and M. C. Villeval (2023). The Way People Lie in Markets: Detectable Vs. Deniable Lies. *Management Science 69*(6), 3340–3357.

Vanberg, C. (2008). Why do people keep their promises? an experimental test of two explanations. *Econometrica 76*(6), 1467–1480.

Weitzel, U. and M. Kirchler (2023). The Banker's Oath And Financial Advice. *Journal of Banking & Finance 148*, 106750.

Zickfeld, J., S. T. Karg, S. S. Engen, A. S. R. Gonzalez, J. Michael, and P. Mitkidis (2022). Committed Dishonesty: A Systematic Meta-Analysis of the Effect of Social Commitment on Dishonest Behavior, doi.org/10.31234/osf.io/j47ng.

ONLINE APPENDIX

A Online Appendix: Literature review

Table A1: Overview of previous experimental studies on oath-taking on various behaviors

Reference	Task	Common knowledge	Compulsory	Findings
Strategic settings				
Jacquemet et al. (2018)	Coordination game	No	No	Positive impact on coordination rates.
Jacquemet et al. (2019)	Deception game	No	No	For selfish lies in loaded frame, the fraction of liars reduced from 35% in No Oath to 16.7% in (voluntary) Oath. Neutral frame does not reduce lies.
Beck (2021)	Deception game	Yes	Yes	Fraction of liars reduced from 42% in No Oath to 20% in (compulsory) Oath. No impact on the size of lies.
Jacquemet et al. (2021)	Deception game	No	No	For selfish lies, fraction of liars goes from 32.8% in No Oath to 14.3% in (voluntary) Oath. No impact on Pareto lies.
Koessler et al. (2021)	Public good game	Yes	Both	Positive impact of both mandatory and voluntary oaths on contributions, but the effect of the voluntary oath fades away more quickly than when it is mandatory.
Davis and Jaber-Lopez (2022)	Binary social dilemma game	No	Both	Impact of both mandatory and voluntary oaths, but crowding-out effect by non oath takers when it is voluntary.
Hergueux et al. (2022)	Public good game	No	No	Positive impact on contributions, according to social types.
Koessler (2022)	Public good game	Yes	No	Positive impact on contributions. More pledges when the majority of the group does so.
Jacquemet et al. (2023)	Trust game	No	No	Oath increases cooperation only with preplay communication. The effect of oath is equivalent to that of a mild fine.
Weitzel and Kirchler (2023)	Audit study	Yes	Yes	Positive impact of reminders of the bankers' oath on real financial advisors' loan recommendations.
Individual decision-making				
Carlsson et al. (2013)	Contingent valuation	-	No	Decrease of the shares of null and very high willingness-to-pay.
Jacquemet et al. (2013)	Second-price auctions	-	No	Decrease of the shares of null and very high willingness-to-pay.
Krüger (2016)	Coin cheating task	-	Yes	Both positive and negative effects depending on gender
Jacquemet et al. (2017)	Voting referenda	-	No	Decrease of the hypothetical bias.
Heinicke et al. (2019)	Number cheating task	-	Yes	Negative impact on extreme lies, no impact on partial lies, in both gain and loss frames.
Koessler et al. (2019)	Tax evasion game	-	No	No impact on compliance except when combined with non-financial rewards.
Schild et al. (2019)	Mind game	-	Yes	Negative impact when lies are unobservable, no impact when lies are observable.
Beck et al. (2020)	Die cheating task	-	Yes	Negative impact on both extreme and partial lies.
Jacquemet et al. (2020)	Tax evasion game	-	No	No impact on extreme lies, negative impact on partial lies.
Kemper et al. (2020)	Discrete choices	-	Yes	Decrease of hypothetical bias in the estimation of willingness-to-pay.
Prima et al. (2020)	Asset reporting	-	Yes	No impact on lies.
Cagala et al. (2021)	Exam-taking	-	Yes	Cheating doubles.
Jacquemet et al. (2021)	Coin cheating task	-	No	Negative impact on extreme lies, no impact on partial lies.
Peer and Feldman (2021)	Reporting perform.	-	Yes	Long-term negative impact on lies.
Akin (2022)	Coin cheating task	-	Yes	Negative effects on undeserved applications.
Babin et al. (2022)	Reporting eye color	-	No	Negative impact on extreme lies, no impact on partial lies.
Babin and Chauhan (2023)	Coin cheating task	-	No	Negative impact on both plausible and implausible misreporting.
Cagala et al. (2023)	Chip cheating task	-	Yes	No impact of commitment requests on lies in both offline and online settings.

B Online Appendix: Instructions [Translated from French]

NO-OATH AND OATH-PRIVATE TREATMENTS

[The printed instructions in Oath-Private treatment are the same as in the No-Oath treatment. However, participants A in this treatment discovered about the oath on their computer screens at the beginning of part 2. During the oath-taking stage, participants A had to type a sentence by which they swear on their honour that during this experiment they will be honest and always tell the truth. Meanwhile, participants B were asked to type a neutral sentence: "The Nobel Prize in Economics for the year 2021 has been awarded to three American economists."]

Welcome *[Common to all treatments]*

Hello and welcome to an experiment on decision-making. Please turn off your phone and put it away. You are not allowed to communicate with other participants during the whole experiment, otherwise you will be excluded from the session and any potential earnings.

During this session you can earn money. The amount of money you will earn depends on your decisions and the decisions of other participants in the session. Please read the instructions carefully. All your decisions and responses will be anonymous.

This session consists of two successive parts. The amount you earn at the end of this session is the sum of your earnings in the different parts plus your participation fee of 5 Euros. During the session we will not talk in Euros but in tokens. The conversion rate from tokens to Euros is as follows:

100 tokens = 4 Euros

You will be paid via bank transfer by the CNRS. To do this, at the end of the session we will ask for your IBAN number. We will also ask you to send us a bank statement (containing the IBAN) in pdf format with the title "your first name-your last name" to the following e-mail address: gatelab[at]gate.cnrs.fr.
Due to administrative and banking delays, the transfer may take up to approximately two weeks.
We promise that your bank details will only be used to make the transfer. This information will be kept separate from the files containing your decisions and the researchers who process your decisions will not have access to your bank details.

You will now find out the instructions for Part 1. The instructions for Part 2 will be explained at the end of Part 1.

Part 1 *[Common to all treatments]*

This part consists of 10 periods. The first 8 periods are training rounds and nothing you decide during these 8 periods counts towards determining your actual earnings. For these 8 training periods the earnings stated are therefore hypothetical.
The only periods that can count towards your earnings in this part are the 9th and 10th periods. At the end of the session, the program will draw period 9 or period 10 and the earnings in the drawn period will constitute your earnings for this part. Each of these two periods has 50 chances out of 100 to be drawn.

Description of the task
In each period, you receive an initial endowment of 100 tokens and you must decide whether to keep these tokens as your earnings for the period or to invest them all in an investment project. If this project is successful, you earn 3

times the number of tokens invested, that is 300 tokens. If the project is not successful, you earn 30 tokens.

Description of the investment project

In each period, 3 cards appear on your screen, face down. Each card can have a star ($\star$) or be blank. Each card has 50 chances out of 100 to have a star and 50 chances out of 100 to be blank. These chances are independent for each card.

Thus, your three cards can have a total of zero stars (which happens with 12.5 chances out of 100), a total of 1 star (which happens with 37.5 chances out of 100), a total of 2 stars (which happens with 37.5 chances out of 100), or a total of 3 stars (which happens with 12.5 chances out of 100).

You must press the "Reveal" button to reveal your three cards. In each case, the cards appear in the following format:

After revealing the cards, you must choose whether or not to invest your 100 tokens.

After your choice, the program draws one of your three cards, each card having the same chance of being drawn (so each card has 1 chance out of 3 to be selected).

- If the card drawn has a star and you have invested, the project is a success and you earn 300 tokens (*i.e.,* endowment of 100 – investment of 100 + earnings of 300).
- If the card drawn is blank and you have invested, the project is not a success and you earn 30 tokens (*i.e.,* endowment of 100 – investment of 100 + earnings of 30).
- If you did not invest your tokens, you keep your initial endowment of 100 tokens and therefore earn 100 tokens.

Whatever your choice is, you are informed at the end of the period whether the card drawn by the program from the three cards has a star or not.

Remember that in these 8 training periods, these earnings are hypothetical and nothing you decide in these periods counts towards determining your actual earnings.

To make sure that you have encountered all possible cases, we have in advance chosen cases that correspond to the probabilities announced in the task description.

Period 9

If drawn at the end of the session, this period determines your actual earnings for this part. The rules and task are the same as in the previous 8 periods. The only difference is in the way you must make your investment choice.

Your screen will show 3 cards, face down. Each card has 50 chances out of 100 to have a star. In this period, you have to make a decision in each of the following 4 scenarios. Would you invest in the project or not if the program announced that among the three cards there are :

Scenario a) 0 stars?

Scenario b) 1 star?

Scenario c) 2 stars?

Scenario d) 3 stars?

Once you have answered these questions, the program will inform you of the total number of stars among your three cards.

Your earnings
If this period 9 is drawn, your earnings are determined by the answer to the scenario that applies. That is, the one that corresponds to the total number of stars among your three cards. For example, suppose that the three cards hide a total of two stars; in this case, your decision in scenario (c) applies. Another example, suppose the three cards hide a total of three stars; in this case, your decision in scenario (d) applies.

The program then draws one of your three cards at random.

- If you have invested in the project and the card drawn has a star, then you earn 300 tokens (endowment of 100 - investment of 100 + earnings of 300).
- If you have invested in the project and the card drawn is blank, then you win 30 tokens (endowment of 100 - investment of 100 + earnings of 30).
- If you have not invested, you earn the 100 tokens of your initial endowment.

As you can see, the principle is the same as in the 8 training periods, but here you make a decision in each possible scenario.
Since only one of your responses will count towards your earnings if this period is drawn, when you make your decision in each scenario it is in your best interest to treat each scenario as if it were the one that actually counts towards your earnings for this part.

Period 10
If drawn at the end of the session, this period determines your actual earnings for this part. As in the previous periods your screen will show 3 cards, face down. Each card has 50 chances out of 100 to have a star. You must again decide whether or not to invest in the project. However, unlike the previous periods, you only have to make one decision without being informed of the number of cards with a star. Only at the end of the session will you be informed by the programme of the total number of stars among your three cards if this period is selected for payment.
Once you have made your decision, the program will then draw one of the three cards at random.

Your earnings
If period 10 is drawn for payment, your earnings are determined as follows:

- If you have invested and the card drawn has a star, then you earn 300 tokens.
- If you have invested and the card drawn is blank, then you earn 30 tokens.
- If you did not invest, you earn the 100 tokens from your initial endowment.

You will be informed of the period drawn (9 or 10), the card drawn and your earnings in that round at the end of the session.

Please read these instructions again. If you have any questions, please raise your hand or press the red button on the side of your desk. We will come and answer your questions in private immediately.

Part 2 *(Instructions distributed after completion of part 1)*

In this part, each of you will be given a role, either 'A' or 'B'. Half of the participants have a role of A and the other half have a role of B. Your role remains the same for the entirety of part 2: you will never change it.
Part 2 has a minimum of 10 periods and a maximum of 30 periods. The exact number of periods has been decided before the start of the session.

In each period, each of you is paired such that there is one participant A and one participant B in each pair. You will never know the identity of the participant you are paired with. At the beginning of each period, you are randomly re-matched with a new
It is unlikely that you will be paired with the same participant two periods in a row.

Your task in each period
Participant A: Participant A sees three cards on his/her screen, face down. Each card can have a star ($\star$) or be blank. Each card has 50 chances out of 100 to have a star and 50 chances out of 100 to be blank. These chances are independent for each card. Thus, participant A can have *a total of* 0 stars, 1 star, 2 stars or 3 stars. Participant A can see how many stars he/she has by pressing the "Reveal" button.

Participant A's task is to then announce his/her total number of stars to participant B with whom he/she is matched with. Participant B cannot see Participant A's cards at any time.

Participant B: Participant B receives an initial endowment of 100 tokens and he/she is only informed of the announcement of Participant A with whom he/she is paired about his/her number of stars. Participant B has to decide whether he/she wants to keep his/her tokens or invest them in Participant A's project.

Determination of earnings
If you are Participant A: You earn a fixed amount of 30 tokens, plus 200 tokens if Participant B has invested in your project.
If you are Participant B: Once you have made your investment decision, the program draws one of Participant A's three cards at random.

- If you have invested in the project and the card drawn has a star, the project is a success and you earn 300 tokens (*i.e.,* endowment of 100 - investment of 100 + earnings of 300).
- If you have invested in the project and the card drawn is blank, the project is not a success and you earn 30 tokens (*i.e.,* endowment of 100 - investment of 100 + earnings of 30).
- If you did not invest, you earn the 100 tokens of your initial endowment.

Whatever your choice is, you are informed at the end of the period whether the card drawn by the program among the three cards has a star or not.

Your screen
At the end of each period, your screen will show a history of what happened in the previous periods. Specifically, you will see four types of information:

1. your announcements in previous periods, if you are Participant A; or the announcements of the different Participants A you were paired with, if you are Participant B;
2. whether the card drawn in previous periods had a star or not, regardless of your role and decision;
3. whether you invested in the previous periods, if you are Participant B; or whether the Participants B you were matched with invested or not, if you are Participant A.
4. your potential earnings in each previous period.

For each of you, **only one period has already been drawn by the program for payment and it is your decision in that period that will determine your earnings in this part**. Thus, it is in your best interest to make your decision in each period as if it were the period that counts towards your earnings in this part.
Remember that **the pairs are re-matched in each period**.

[***ADDITIONAL PARAGRAPH IN OATH-CK TREATMENT ONLY***]

Oath
At the beginning of part 2, **Participants A will have to take an oath about the truthfulness of their announcements in part 2**. By this oath, they will swear on their honour that during this experiment they will be honest and always tell the truth.

[After the instructions and role assignment, participants A in Oath-CK treatment swore the oath by typing it on their computer. Participants B did not type anything during this stage. After all participants A took the oath, every participant was informed that the oath had been taken.]

- - -

Please read these instructions again. If you have any questions, please raise your hand or press the red button. We will come and answer your questions in private immediately.

- - -

End of the instructions in part 2 *[Displayed on computer screen only]*

[At the end of part 2 but before informing participants of their final gains, we elicited beliefs, depending on their role in the Announcement Game. For Participants B (investors), we asked the following question about their beliefs concerning the honesty of advisors in their session.]
Please estimate below the percentage of true announcements among all the announcements you received in part 2. Enter integer values between 0 and 100. At the end of the session the program will draw one of these questions. If your answer is correct, plus or minus 5%, you will earn an extra 0.50€.

In your opinion, what percentage of announcements were true when Participant A announced:

- 0 stars: _%
- 1 star: _%
- 2 stars:_%

- 3 stars:_%

[For Participants A (advisors), we asked two questions (in randomized order at an individual level). One elicited their beliefs about the honesty of advisors in their session, which reads as follows:]
Please estimate below the percentage of true announcements among all announcements made in part 2 by all participants A in this session. Enter integer values between 0 and 100. At the end of the session the program will draw one of these questions. If your answer is correct, plus or minus 5%, you will earn an extra 0.50€.

In your opinion, what percentage of announcements were true when participants A in this session announced:

- 0 stars: _%
- 1 star: _%
- 2 stars:_%
- 3 stars:_%

[The other question asked about their beliefs about what investors think about their honesty, which reads as follows:]
Your participant B is answering the following question: "In your opinion, what percentage of announcements were true when participant A told you?".
What do you think your participant B's answer was to these questions?
Enter integer values between 0 and 100. At the end of the session the program will draw one of these questions. If your answer is correct, plus or minus 5%, you win an extra 0.50€.

- 0 stars: _%
- 1 star: _%
- 2 stars:_%
- 3 stars:_%

[After the belief elicitations, participants were informed about their final gains (flat payment and additional payoffs from part 1, part 2, belief questions). They then completed a final questionnaire consisting of the following information:]

- Age
- Gender
- Status
- School
- *[Advisors only]* In part 2, if sometimes you did not announce the correct number of stars or if you announce the correct number of stars all the time, can you explain why you did so? If you have different reasons, can you rank them?
- *[Investors only]* In part 2, if sometimes you have invested or not in Participant A's project, can you explain why? If you have different reasons, can you rank them.
- What do you think this experiment is testing?
- If you had to leave a piece of advice for a successor in your role, what would it be?

\- - -

ASYM-ASYNC TREATMENT

[The written instructions for part 1 of the experiment are the same as in the other treatments. The following instruction was given to participants for part 2.]

Part 2

Some weeks ago, participants from the same subjects pool as you took part in an experimental session. These participants performed the same task as you in part 1, under exactly the same conditions as you did. Please read below an extract from the instructions they were given for part 2 (text in italics). The supplementary instructions that only concern you will be given after.

Extract from the instructions given to past participants

In this part, each of you will be given a role, either 'A' or 'B'. Half of the participants have a role of A and the other half have a role of B. Your role remains the same for the entirety of part 2: you will never change it.
Part 2 has a minimum of 10 periods and a maximum of 30 periods. The exact number of periods has been decided before the start of the session.

In each period, each of you is paired such that there is one participant A and one participant B in each pair. You will never know the identity of the participant you are paired with. At the beginning of each period, you are randomly re-matched with a new p
It is unlikely that you will be paired with the same participant two periods in a row.

Your task in each period
Participant A: Participant A sees three cards on his/her screen, face down. Each card can have a star (⋆) or be blank. Each card has 50 chances out of 100 to have a star and 50 chances out of 100 to be blank. These chances are independent for each card. Thus, participant A can have a total of *0 stars, 1 star, 2 stars or 3 stars. Participant A can see how many stars he/she has by pressing the "Reveal" button.*

Participant A's task is to then announce his/her total number of stars to participant B with whom he/she is matched with. Participant B cannot see Participant A's cards at any time.

Participant B: Participant B receives an initial endowment of 100 tokens and he/she is only informed of the announcement of Participant A with whom he/she is paired about his/her number of stars. Participant B has to decide whether he/she wants to keep his/her tokens or invest them in Participant A's project.

Determination of earnings
If you are Participant A: You earn a fixed amount of 30 tokens, plus 200 tokens if Participant B has invested in your project.

If you are Participant B: Once you have made your investment decision, the program draws one of Participant A's three cards at random.

- *If you have invested in the project and the card drawn has a star, the project is a success and you earn 300 tokens (i.e., endowment of 100 - investment of 100 + earnings of 300).*
- *If you have invested in the project and the card drawn is blank, the project is not a success and you earn 30 tokens (i.e., endowment of 100 - investment of 100 + earnings of 30).*

- *If you did not invest, you earn the 100 tokens of your initial endowment.*

Whatever your choice is, you are informed at the end of the period whether the card drawn by the program among the three cards has a star or not.

Your screen
At the end of each period, your screen will show a history of what happened in the previous periods. Specifically, you will see four types of information:

1. *Your announcements in previous periods, if you are Participant A; or the announcements of the different Participants A you were paired with, if you are Participant B;*

2. *Whether the card drawn in previous periods had a star or not, regardless of your role and decision;*

3. *Whether you invested in the previous periods, if you are Participant B; or whether the Participants B you were matched with invested or not, if you are Participant A.*

4. *Your potential earnings in each previous period.*

For each of you, ***only one period has already been drawn by the program for payment and it is your decision in that period that will determine your earnings in this part****. Thus, it is in your best interest to make your decision in each period as if it were the period that counts towards your earnings in this part.*

Remember that ***the pairs are re-matched in each period****.*

End of the extract

Supplementary instructions for participants in today's session

The previous instructions apply to you in this part as well but there are some differences from the previous extract:

1. **All participants in today's session are assigned to <u>the role B</u>**; you will keep this role throughout the part.

2. Participants in **role A are not present in this session**: they have already participated in a previous session with other participants in role B. The excerpt of the instructions you read above was for these past participants.

3. At the beginning of each period, you will be matched with a participant A from a previous session randomly selected by the program.

4. Before you decide whether to invest, you will be informed of the announcement made by this participant A from a previous session after the initial draw of his/her three cards.

5. After deciding whether or not to invest, you will be informed of the card that was randomly drawn by the program from the three cards of this participant A from a previous session. This card drawn will determine your earnings for the period if you have decided to invest.

6. Your investment decisions will only affect your earnings and not the earnings of the participants A from previous sessions with whom you are matched.

Summary

This part has a minimum of 10 periods and a maximum of 30 periods. The exact number of periods has been decided before the start of the session.

In each period you will be informed about the number of stars (0, 1, 2 or 3) announced by a participant A from a previous session in a period drawn at random. You will then have to decide whether or not to invest in the project of participant A from the previous session.

- If you have invested in the project and the card drawn is a star, the project is a success and you earn 300 tokens.
- If you have invested in the project and the card drawn is blank, the project is not a success and you earn 30 tokens.
- If you have not invested, you earn the 100 tokens of your initial endowment.

Whatever your choice is, you are informed at the end of the period whether the card drawn by the program among the three cards has a star or not.

For each of you, **only one period has already been drawn by the program for payment and it is your decision in that period that will determine your earnings in this part**. Thus, it is in your best interest to make your decision in each period as if it were the period that counts towards your earnings in this part.

Remember that **in each period the program randomly selects a decision** (*i.e.,* an announcement made after an initial draw of three cards) **from all the decisions made by participants A in previous sessions. You are re-matched with a new participant A in each period** and it is very unlikely that the program will select the same participant A for you two periods in a row.

- - -

Please read these instructions again. If you have any questions, please raise your hand or press the red button. We will come and answer your questions immediately in private.

- - -

End of the instructions in part 2 *[Displayed on computer screen only]*

[In the Asym-Async treatment, at the end of part 2 but before informing participants of their final gains, we elicited beliefs, depending on their role in the Announcement Game like in the other treatments. However, after the belief elicitation about the proportion of honest announcements, participants in this treatment, who all had the role of investors, were asked to compare their responses to their initial expectations about advisors' honesty before starting the Announcement Game.]

We remind you below of your estimates of the proportions of true announcements received in part 2. Please let us know how well these estimates match your initial expectations before you began part 2. Specifically, indicate for each scenario whether your estimates are:

- More or less equal to your initial expectations

- Higher than your initial expectations (before the part started, you expected a lower proportion of true announcements)
- Lower than your initial expectations (before the part started, you expected a higher proportion of true announcements)

Your estimate of the proportion of truthful announcements are ...

- 0/1/2/3 stars (corresponding estimate shown here):
  - More or less equal your initial expectations
  - Higher than your initial expectations ...
  - Lower than your initial expectations ...

- - -

SYM-ASYNC TREATMENT

[The written instructions for part 1 of the experiment are the same as in the other treatments. In part 2, like in the Asym-Async treatment, participants were first given the extract of the instructions used in Oath-CK treatment. Note that the extract part for the Sym-Async treatment included information about the oath, as in the Oath-CK treatment.]

Part 2

Some weeks ago, participants from the same subjects pool as you took part in an experimental session. These participants performed the same task as you in part 1, under exactly the same conditions as you did. Please read below an extract from the instructions they were given for part 2 (text in italics). The supplementary instructions that only concern you will be given after.

Extract from the instructions given to past participants

In this part, each of you will be given a role, either 'A' or 'B'. Half of the participants have a role of A and the other half have a role of B. Your role remains the same for the entirety of part 2: you will never change it.
Part 2 has a minimum of 10 periods and a maximum of 30 periods. The exact number of periods has been decided before the start of the session.

In each period, each of you is paired such that there is one participant A and one participant B in each pair. You will never know the identity of the participant you are paired with. At the beginning of each period, you are randomly re-matched with a new p It is unlikely that you will be paired with the same participant two periods in a row.

Your task in each period
Participant A: Participant A sees three cards on his/her screen, face down. Each card can have a star (⋆) or be blank. Each card has 50 chances out of 100 to have a star and 50 chances out of 100 to be blank. These chances are independent for each card. Thus, participant A can have a total of *0 stars, 1 star, 2 stars or 3 stars. Participant A can see how many stars he/she has by pressing the "Reveal" button.*

Participant A's task is to then announce his/her total number of stars to participant B with whom he/she is matched with. Participant B cannot see Participant A's cards at any time.

Participant B: Participant B receives an initial endowment of 100 tokens and he/she is only informed of the announcement of Participant A with whom he/she is paired about his/her number of stars. Participant B has to decide whether he/she wants to keep his/her tokens or invest them in Participant A's project.

Determination of earnings
If you are Participant A: You earn a fixed amount of 30 tokens, plus 200 tokens if Participant B has invested in your project.

If you are Participant B: Once you have made your investment decision, the program draws one of Participant A's three cards at random.

- *If you have invested in the project and the card drawn has a star, the project is a success and you earn 300 tokens (i.e., endowment of 100 - investment of 100 + earnings of 300).*
- *If you have invested in the project and the card drawn is blank, the project is not a success and you earn 30 tokens (i.e., endowment of 100 - investment of 100 + earnings of 30).*

- *If you did not invest, you earn the 100 tokens of your initial endowment.*

Whatever your choice is, you are informed at the end of the period whether the card drawn by the program among the three cards has a star or not.

Your screen
At the end of each period, your screen will show a history of what happened in the previous periods. Specifically, you will see four types of information:

1. *Your announcements in previous periods, if you are Participant A; or the announcements of the different Participants A you were paired with, if you are Participant B;*

2. *Whether the card drawn in previous periods had a star or not, regardless of your role and decision;*

3. *Whether you invested in the previous periods, if you are Participant B; or whether the Participants B you were matched with invested or not, if you are Participant A.*

4. *Your potential earnings in each previous period.*

For each of you, ***only one period has already been drawn by the program for payment and it is your decision in that period that will determine your earnings in this part****. Thus, it is in your best interest to make your decision in each period as if it were the period that counts towards your earnings in this part.*

Remember that ***the pairs are re-matched in each period****.*

Oath
At the beginning of part 2, ***Participants A will have to take an oath about the truthfulness of their announcements in part 2****. By this oath, they will swear on their honour that during this experiment they will be honest and always tell the truth.*

End of the extract

Supplementary instructions for participants in today's session

The previous instructions apply to you in this part as well but there are some differences from the previous extract:

1. **All participants in today's session are assigned to <u>the role B</u>**; you will keep this role throughout the part.

2. Participants in **role A are not present in this session**: they have already participated in a previous session with other participants in role B. The excerpt of the instructions you read above was for these past participants. Before starting the part, **all participants in role A had to take an oath on honor** that during the experiment they would behave honestly and would always tell the truth. These participants in role A had to type the text of the oath on their computer and this was made known to the participants in role B.

3. **At the beginning of each period, you will be matched with a participant A, randomly selected from a subset of these participants A.**

4. Before you decide whether to invest, you will be informed of the announcement made by this participant A from a previous session after the initial draw of his/her three cards.

5. After deciding whether or not to invest, you will be informed of the card that was randomly drawn by the program from the three cards of this participant A from a previous session. This card drawn will determine your earnings for the period if you have decided to invest.

6. Your investment decisions will only affect your earnings and not the earnings of the participants A from previous sessions with whom you are matched.

Summary
This part has a minimum of 10 periods and a maximum of 30 periods. The exact number of periods has been decided before the start of the session.

In each period you will be informed about the number of stars (0, 1, 2 or 3) announced by a participant A from a previous session in a period drawn at random. You will then have to decide whether or not to invest in the project of participant A from the previous session.

- If you have invested in the project and the card drawn is a star, the project is a success and you earn 300 tokens.
- If you have invested in the project and the card drawn is blank, the project is not a success and you earn 30 tokens.
- If you have not invested, you earn the 100 tokens of your initial endowment.

Whatever your choice is, you are informed at the end of the period whether the card drawn by the program among the three cards has a star or not.

For each of you, **only one period has already been drawn by the program for payment and it is your decision in that period that will determine your earnings in this part**. Thus, it is in your best interest to make your decision in each period as if it were the period that counts towards your earnings in this part.

Remember that **in each period you will face a decision** (i.e., an announcement made after an initial draw of three cards) **made by one participant A, chosen at random from a subset of the participants A from previous sessions. You are rematched to a new participant A each period** and it is unlikely that the program will select the same participant A for you two periods in a row.

- - -

Please read these instructions again. If you have any questions, please raise your hand or press the red button. We will come and answer your questions immediately in private.

- - -

End of the instructions in part 2 *[Displayed on computer screen only]*

[In the Sym-Async treatment, at the end of part 2 but before informing participants of their final gains, we elicited beliefs, depending on their role in the Announcement Game like in the other treatments. After the belief elicitation about the proportion of honest announcements, participants in this treatment were asked to compare their responses to their initial expectations about advisors' honesty before starting part 2. Note that the wording of the question was slightly adapted between Asym-Async and Sym-Async treatments.]

We will ask you questions about the possible differences between your initial expectations (after reading the instructions but before starting the part) and your experience during the experiment regarding the reliability of the announcements received from participants A.

- When the announcement was 0 stars:
  - My experience has shown me that these announcements were LESS RELIABLE than my expectations (the 0-star announcements were less truthful than expected)
  - My experience has shown me that these announcements were MORE RELIABLE than I expected (the 0 announcements were more truthful than expected)
  - My experience has shown me that the reliability of these announcements was NEARLY EQUAL to my expectations
- When the announcement was 1 star:
  - My experience has shown me that these announcements were LESS RELIABLE than my expectations (the 1-star announcements were less truthful than expected)
  - My experience has shown me that these announcements were MORE RELIABLE than I expected (the 1-star announcements were more truthful than expected)
  - My experience has shown me that the reliability of these announcements was NEARLY EQUAL to my expectations
- When the announcement was 2 stars:
  - My experience has shown me that these announcements were LESS RELIABLE than my expectations (the 2-star announcements were less truthful than expected)
  - My experience has shown me that these announcements were MORE RELIABLE than I expected (the 2-star announcements were more truthful than expected)
  - My experience has shown me that the reliability of these announcements was NEARLY EQUAL to my expectations
- When the announcement was 3 stars:
  - My experience has shown me that these announcements were LESS RELIABLE than my expectations (the 3-star announcements were less truthful than expected)
  - My experience has shown me that these announcements were MORE RELIABLE than I expected (the 3-star announcements were more truthful than expected)
  - My experience has shown me that the reliability of these announcements was NEARLY EQUAL to my expectations

- - -

C Online Appendix: Tables

Tables C4 and C5 present the results of random-effects Linear Probability models (GLS) in which the dependent variable is the decision to make a detectable or a deniable lie, respectively. Models (1) and (2) consider all the advisors. Models (3) and (4) are restricted to the advisors who lied at least once in part 2.

Alternatively to models (3) and (4), we estimated two-step Heckman models, estimating in a first step the probability to make a lie and, in the second step, the probability to make a detectable or a deniable lie, using gender as the identifying variable. However, since the IMR was not significant, meaning that there is no need for correcting for a selection bias, we omit these regressions and only report the GLS models.

Table C1: Summary statistics of advisors

	(1) No-Oath Mean	SD	(2) Oath-CK Mean	SD	(3) Oath-Private Mean	SD	(1-2)	(1-3) *p-value*	(2-3)
Age (Years)	21.34	3.62	20.90	1.84	20.92	2.13	0.626	0.656	0.968
Male (%)	0.66	0.48	0.44	0.50	0.44	0.50	0.007***	0.007***	0.969
Business school (%)	0.50	0.50	0.61	0.49	0.40	0.49	0.170	0.208	0.009***
Number of observations	76		77		73				

Notes: This table summarizes the socio-demographic characteristics of participants in the role of advisors for each treatment. The *p-value* reported using ranksum test for age and tests of proportion for gender (coded 1 as male, 0 otherwise) and business school (coded 1, 0 otherwise). Regression analyzes controlled for any effects of these characteristics.

Table C2: Summary statistics of investors

	(1) No-Oath Mean	SD	(2) Oath-CK Mean	SD	(3) Oath-Private Mean	SD	(4) Asym-Async Mean	SD	(5) Sym-Async Mean	SD
Age (Years)	20.82	2.03	20.60	1.32	21.42	2.11	21.71	1.85	20.99	1.84
Male (%)	0.61	0.49	0.38	0.49	0.49	0.50	0.44	0.50	0.64	0.48
Business school (%)	0.54	0.50	0.66	0.48	0.48	0.50	0.48	0.50	0.36	0.48
Number of observations	76		77		73		77		78	

Notes: This table summarizes the socio-demographic characteristics of participants in the role of investors for each treatment. The *p-values* are reported in the next table.

Table C3: Between-treatment comparisons of socio-demographics of investors

	No-Oath *vs.* Oath-CK	No-Oath *vs.* Oath-Private	Oath-CK *vs.* Oath-Private	Oath-CK *vs.* Asym-Async	Oath-Private *vs.* Asym-Async	Oath-CK *vs.* Sym-Async	No-Oath *vs.* Sym-Async
Age (Years)	$p = 0.935$	$p = 0.024$	$p = 0.007$	$p < 0.001$	$p = 0.126$	$p = 0.212$	$p = 0.330$
Male (%)	$p = 0.005$	$p = 0.169$	$p = 0.150$	$p = 0.413$	$p = 0.527$	$p = 0.001$	$p = 0.647$
Business school (%)	$p = 0.121$	$p = 0.464$	$p = 0.024$	$p = 0.023$	$p = 0.990$	$p < 0.001$	$p = 0.024$

Note: This table reports the p-values from ranksum tests (for age) and tests of proportions (for gender and school) comparing the summary statistics of investors across treatments. Regression analyzes controlled for any effects of these characteristics.

Table C4: Probability of making detectable lies

Dependent variable:	All advisors		Lied at least once	
Detectable lie	(1)	(2)	(3)	(4)
No-Oath	*Ref.*	*Ref.*	*Ref.*	*Ref.*
Oath-CK	-0.201***	-0.210***	-0.121*	-0.166**
	(0.038)	(0.049)	(0.071)	(0.080)
Oath-Private	-0.147***	-0.153***	-0.112**	-0.125**
	(0.040)	(0.051)	(0.052)	(0.062)
Period	0.001	0.001	0.002	0.001
	(0.001)	(0.004)	(0.002)	(0.004)
Male	0.064**	0.064**	0.058	0.058
	(0.030)	(0.030)	(0.048)	(0.048)
Age	-0.008**	-0.008**	-0.010**	-0.010**
	(0.003)	(0.003)	(0.004)	(0.004)
Business school	-0.015	-0.015	-0.008	-0.008
	(0.030)	(0.030)	(0.047)	(0.047)
Oath-CK X Period	-	0.001	-	0.005
		(0.004)		(0.005)
Oath-Private X Period	-	0.001	-	0.001
		(0.004)		(0.005)
Constant	0.382***	0.387***	0.430***	0.441***
	(0.088)	(0.094)	(0.108)	(0.113)
Number of observations	3551	3551	2096	2096
Number of clusters	226	226	133	133

Notes: This table presents the coefficients from random-effects Linear Probability models (GLS) in which the dependent variable is the decision to make a detectable lie (*i.e.*, announcing three stars, conditional on observing less than three stars), for all advisors (models (1) and (2)) and those who lied at least once (models (3) and (4)). Independent variables in model (1) include treatment dummies (with No-Oath treatment as the reference category), period, a male dummy, age (in years), and a dummy for being student at the business school. Model (2) includes interaction terms between the treatment dummies and the period variable. Models (3) and (4) correspond to models (1) and (2) respectively for the subset of advisors who lied at least once. Standard errors are clustered at the individual level.

Table C5: Probability of making deniable lies

Dependent variable:	All advisors		Lied at least once	
Deniable lie	(1)	(2)	(3)	(4)
No-Oath	*Ref.*	*Ref.*	*Ref.*	*Ref.*
Oath-CK	-0.406***	-0.424***	-0.094	-0.252**
	(0.055)	(0.065)	(0.086)	(0.121)
Oath-Private	-0.171***	-0.242***	0.013	-0.135
	(0.063)	(0.072)	(0.061)	(0.086)
Period	0.007***	0.004	0.012***	0.004
	(0.002)	(0.004)	(0.003)	(0.005)
Male	0.109**	0.109**	0.034	0.036
	(0.049)	(0.049)	(0.058)	(0.057)
Age	0.008	0.008	0.007	0.007
	(0.007)	(0.007)	(0.006)	(0.006)
Business school	-0.007	-0.007	0.031	0.034
	(0.047)	(0.047)	(0.056)	(0.056)
Oath-CK X Period	-	0.002	-	0.017*
		(0.005)		(0.009)
Oath-Private X Period	-	0.007	-	0.016**
		(0.005)		(0.007)
Constant	0.277*	0.307*	0.325**	0.401**
	(0.160)	(0.163)	(0.153)	(0.157)
Number of observations	2039	2039	1179	1179
Number of clusters	226	226	133	133

Notes: This table presents the coefficients from random-effects Linear Probability models (GLS) in which the dependent variable is the decision to make a deniable lie (*i.e.*, announcing one or two stars conditional on observing zero or one star, respectively) for all advisors (models (1) and (2)) and those who lied at least once (models (3) and (4)). Independent variables in model (1) include treatment dummies (with No-Oath treatment as the reference category), period, a male dummy, age (in years), and a dummy for being student at the business school. Model (2) includes interaction terms between the treatment dummies and the period variable. Models (3) and (4) correspond to models (1) and (2) respectively for the subset of advisors who lied at least once. Standard errors are clustered at the individual level.

Table C6: Determinants of the investment decision (two- and three-star announcements)

Dependent variable: Investment decision	(1)	(2)
No-Oath	*Ref.*	*Ref.*
Oath-CK	0.171***	0.096**
	(0.035)	(0.038)
Oath-Private	0.049	-0.001
	(0.040)	(0.041)
Asym-Async	0.166***	0.101**
	(0.039)	(0.041)
Sym-Async	0.071*	0.065
	(0.039)	(0.040)
Period	-0.005***	-0.004***
	(0.001)	(0.001)
Male	0.026	0.030
	(0.024)	(0.024)
Age	0.000	0.001
	(0.006)	(0.006)
Business school	-0.011	-0.017
	(0.025)	(0.025)
Index of detection	-	-0.685***
		(0.107)
% of blank cards on 2 stars	-	-0.064**
		(0.027)
Constant	0.755***	0.835***
	(0.131)	(0.132)
Number of observations	6146	5552
Number of clusters	381	381
Wald tests	*p-value*	
Asym-Async *vs.* Oath-CK	0.894	0.893
Sym-Async *vs.* Oath-CK	0.006	0.404

Notes: This table presents the coefficients from random-effects Linear Probability models (GLS) in which the dependent variable is the decision to invest, conditioned on receiving a two- or three-star announcement. Independent variables in model (1) include treatment dummies (with No-Oath treatment as the reference category), period, a male dummy, age (in years), and a dummy for being student at the business school. Model (2) includes controls for the percentage of the time a lie has been detected up to the period ('Index of detection') and the percentage of blank cards on two-star announcements up to the period ('% of blank cards on 2 Stars'). Note that interaction terms between the treatments and these two control variables are not significantly different from zero; thus, we do not report this additional model. Standard errors are clustered at the individual level.

Table C7: Time trends of investment decision by treatment (all announcements)

Dependent variable Investment decision	(1) No-Oath	(2) Oath-CK	(3) Oath-Private	(4) Asym-Async	(5) Sym-Async
Period	-0.006**	-0.000	-0.001	0.001	-0.007***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
Constant	0.718***	0.574***	0.583***	0.540***	0.818***
	(0.032)	(0.027)	(0.034)	(0.028)	(0.027)
Number of observations	1368	1386	1314	2079	2106
Number of clusters	76	77	73	77	78

Notes: This table presents the coefficients from random-effects Linear Probability models (GLS) in which the dependent variable is the decision to invest. The only independent variable is the period. Standard errors are clustered at the individual level.

D Online Appendix: Beliefs

We elicited the beliefs of participants at the end of the sessions. Investors were asked to estimate the proportion of truthful announcements for each number of stars made to them in part 2 (*i.e.,* their first-order beliefs about the honesty of advisors). The following tables report descriptive statistics and non-parametric tests on these beliefs.

Summary of belief elicitation data Belief data is notoriously noisy, and this portion of our data is no exception. We note that our belief data are nonetheless often qualitatively consistent with the actions that we observe among our subjects. For example, even though the point estimates do not match actual investment rates, we see in Table D1 that in the No-Oath, Oath-Private and Sym-Async treatments subjects are more distrustful when they are faced with a higher announcement than when they are faced with a lower announcement (and indeed in these treatments advisors are in large part dishonest). In contrast, in the Oath-CK and Asym-Async treatments, which are the treatments in which investors generally experience truthful announcements, the credibility of announcements is more stable accross announcements than in the treatments in which adivsors are dishonest. The data in Table D2 also shows that the cross-treatment comparisons often match the patterns we observe in actual investment rates, for example, with subjects trusting three-star announcements significantly more in Oath-CK than in No-Oath (first column), or that there are no significant differences in how much they trust a three-star announcement across the No-Oath and Sym-Async treatments (last column).

The within-treatment data on the first order beliefs of advisors on the truthfulness of other advisors' announcements are also very noisy (see Table D3). However, the cross-treatment comparisons of these beliefs align with what would be expected given behavior. For example, as Table D4 shows, beliefs on truthfulness levels of zero-star announcements are no different across treatments. In that same table we also see that there are large and significant cross-treatment differences in beliefs about truthfulness of three-star announcements: it is the highest in Oath-CK, followed by Oath-Private and then No-Oath.

Finally, the second-order beliefs of advisors regarding what they think investors believe the proportion of truthful announcements by number of stars is (see Tables D5 and D6) generally follow the same patterns as the first order beliefs of advisors: for example, advisors believe that investors will believe a three-star announcement more in the Oath-CK than in the Oath-Private, and more in the Oath-Private than in the No-Oath treatment.

Table D1: First-order beliefs of investors

Announcements	No Oath	Oath-CK	Asym-Async	Oath-Private	Sym-Async
0 Stars	81.00	86.68	79.25	81.52	77.84
1 Star	82.69	88.30	75.64	78.29	80.66
2 Stars	53.83	80.82	71.51	62.05	60.78
3 Stars	44.26	82.53	74.12	60.64	49.65

Note: This table reports the investors' first-order beliefs about the proportion of truthful announcements, by number of stars announced and by treatment.

Table D2: Between-treatment comparisons of first-order beliefs of investors

	No-Oath *vs.* Oath-CK	No-Oath *vs.* Oath-Private	No-Oath *vs.* Asym-Async	Oath-CK *vs.* Oath-Private	Oath-CK *vs.* Asym-Async	Oath-Private *vs.* Asym-Async	Oath-CK *vs.* Sym-Async	No-Oath *vs.* Sym-Async
0 Stars	$p = 0.716$	$p = 0.684$	$p = 0.979$	$p = 0.917$	$p = 0.614$	$p = 0.532$	$p = 0.773$	$p = 0.998$
1 Star	$p = 0.239$	$p = 0.040$	$p = 0.002$	$p = 0.071$	$p = 0.011$	$p = 0.294$	$p = 0.730$	$p = 0.187$
2 Stars	$p < 0.001$	$p = 0.034$	$p < 0.001$	$p < 0.001$	$p = 0.009$	$p = 0.015$	$p < 0.001$	$p = 0.038$
3 Stars	$p < 0.001$	$p = 0.002$	$p < 0.001$	$p < 0.001$	$p = 0.023$	$p = 0.010$	$p < 0.001$	$p = 0.175$

Note: This table reports the p-values from rank-sum tests comparing the 1st-order beliefs of investors across treatments.

Table D3: First-order beliefs of advisors

Announcements	No-Oath	Oath-CK	Oath-Private
0 Stars	63.62	75.90	69.77
1 Star	63.16	79.30	70.16
2 Stars	53.72	84.55	69.12
3 Stars	50.99	87.91	72.77

Note: This table displays the 1st-order beliefs of advisors about the proportion of truthful announcements by the other advisors, by number of stars announced and by treatment.

Table D4: Between-treatment comparisons of first-order beliefs of advisors

	No-Oath *vs.* Oath-CK	No-Oath *vs.* Oath-Private	Oath-CK *vs.* Oath-Private
0 Stars	$p = 0.124$	$p = 0.281$	$p = 0.690$
1 Star	$p = 0.011$	$p = 0.160$	$p = 0.314$
2 Stars	$p < 0.001$	$p < 0.001$	$p < 0.001$
3 Stars	$p < 0.001$	$p < 0.001$	$p = 0.009$

Notes: This table reports the p-values from rank-sum tests comparing the 1st-order beliefs of advisors across treatments.

Table D5: Second-order beliefs of advisors

Announcements	No-Oath	Oath-CK	Oath-Private
0 Stars	56.07	65.19	53.07
1 Star	56.36	68.21	55.45
2 Stars	57.32	79.71	59.14
3 Stars	48.88	88.16	67.71

Note: This table displays the second-order beliefs of advisors, that is, what they think investors believe about the proportion of truthful announcements, by number of stars announced and by treatment.

Table D6: Between-treatment comparisons of the second-order beliefs of advisors

	No-Oath *vs.* Oath-CK	No-Oath *vs.* Oath-Private	Oath-CK *vs.* Oath-Private
0 Stars	$p = 0.225$	$p = 0.939$	$p = 0.162$
1 Star	$p = 0.095$	$p = 0.784$	$p = 0.046$
2 Stars	$p < 0.001$	$p = 0.603$	$p < 0.001$
3 Stars	$p < 0.001$	$p = 0.001$	$p < 0.001$

Note: This table reports the p-values from rank-sum tests comparing the 2nd-order beliefs of advisors across treatments.

E Online Appendix: Figures

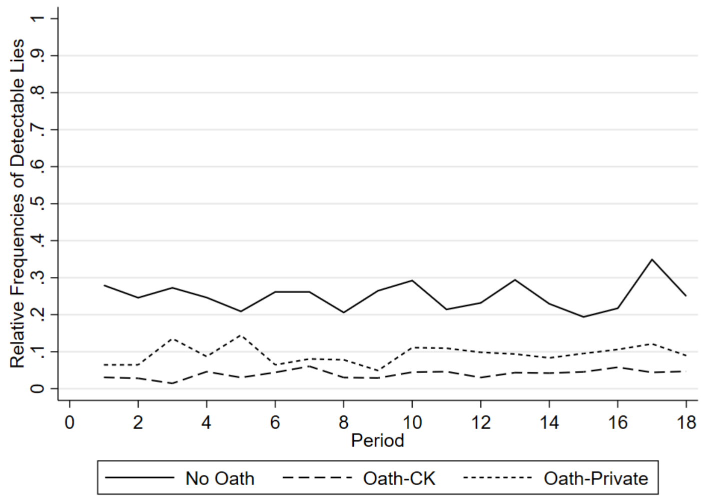

Figure E1: Evolution of the relative frequencies of detectable lies, by period and treatment

Note: The figure displays the evolution of the relative frequency of detectable lies over time, by treatment, including only the observations in which these lies were possible, that is, excluding the observations in which three stars were observed.

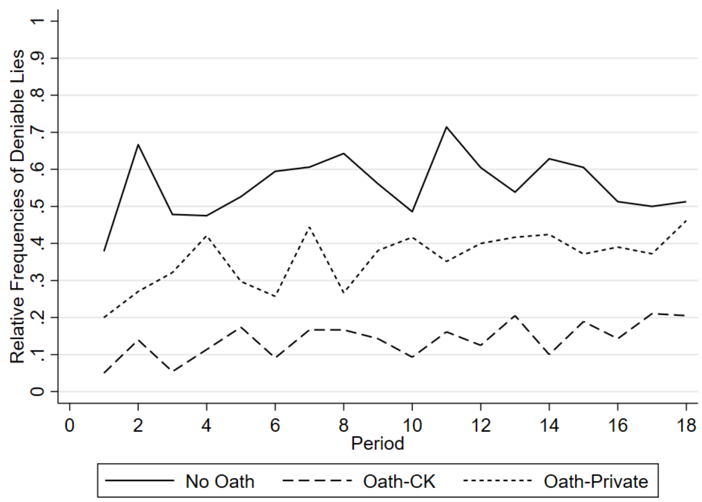

Figure E2: Evolution of the relative frequencies of deniable lies, by period and treatment

Note: The figure displays the evolution of the relative frequency of deniable lies over time, by treatment, including only the observations in which these lies were possible, that is, excluding the observations in which two or three stars were observed.

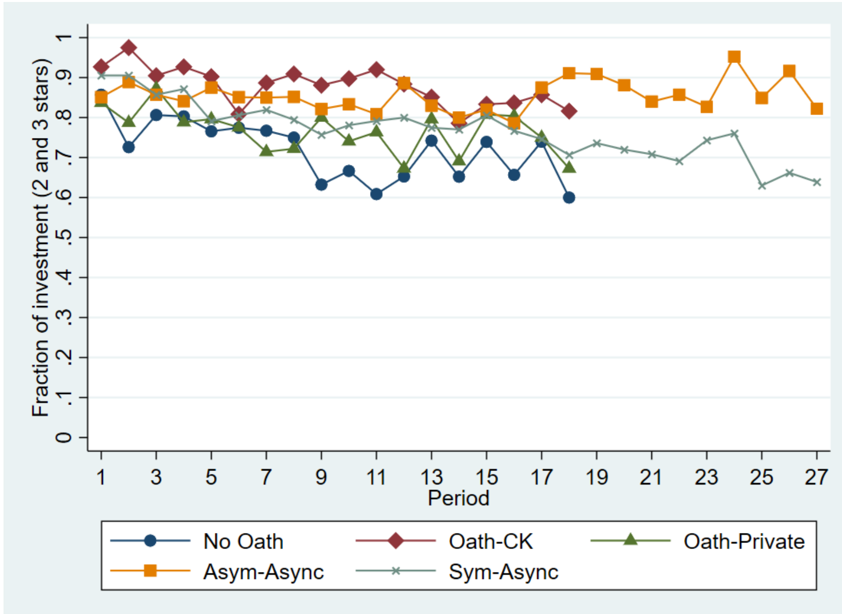

Figure E3: Evolution of the fraction of investment on 2 and 3 stars, by period and treatment

Note: The figure displays the evolution of the fraction of investments on two- and three-star announcements over time, by treatment. There were 18 periods in the No-Oath, Oath-CK, and Oath-Private treatments, and 27 periods in the Asym-Async and Sym-Async treatments.

F Online Appendix: Survey on Prolific

We report in subsection F.1 the questionnaire used in the survey conducted in December 2022 via Prolific with 395 individuals residing in the Netherlands (198) and France (197). The survey was collected via Google Forms in Dutch and French respectively, and translated back to English. We analyze the responses in subsection F.2.

F.1 Questionnaire

Consent
This study includes a short survey conducted by Chloe Tergiman (Pennsylvania State University), Marie Claire Villeval (University of Lyon), and Sorravich Kingsuwankul (University of Lyon). It takes about 4 minutes to complete the survey.
Participation is voluntary. You can withdraw from the survey at any time without giving a reason. This will not affect the relationship between you, the researchers or Prolific in any way. All aspects of the study are confidential and anonymous. We will not ask for your name or any information that may identify you during the study. A report of this survey may be submitted for publication, but all information will be used in an aggregate form.
Upon completion of this survey, you will receive a fixed fee of 1€. If you have specific questions about the survey, you can contact us by e-mail at kingsuwankul[at]gate.cnrs.fr
If you agree with the above terms, please select "I agree" to continue with the survey.

- I agree *(Continue)*

- I do not agree *(Exit survey)*

Please enter your Prolific ID:
(Input box here)

Do you live in the Netherlands/France?

- Yes *(Continue)*
- No *(Exit survey)*

Q1. Could you say in general that most people can be trusted or that you have to be very careful when dealing with people?

- Most people can be trusted.
- You have to be very careful.

Q2. Please indicate on a 5-point scale from 1 "I don't trust at all" to 5 "I trust completely", how much you trust each of these groups:

- Doctors
- Members of the parliament
- Judges
- Firefighters
- Bankers

Q3A. *(If reported trust in bankers in Q2 is low: 1-3)* In your response to the previous question, you have indicated that **bankers cannot be completely trusted**. Can you explain the reason(s)? You can choose up to three reasons.

- Banking culture is not trustworthy.
- I/someone I know had a bad experience with bankers/banks.
- The banking industry attracts more dishonest employees than other sectors.
- Bankers' unethical behavior is difficult to detect.
- Unethical behavior by employees is usually not punished by the banks and can sometimes even be rewarded by bonuses.
- Bankers do not prioritize the customers' interests over those of the bank.

Is there another reason, not mentioned above, why you think bankers cannot be completely trusted? If not, you can skip this question.

Q3B. *(If reported trust in bankers in Q2 is high: 4 or 5)* In your response to the previous question, you have indicated that **bankers can generally be trusted**. Can you explain the reason(s)? You can choose up to three reasons.

- Scandals are rare and don't reflect the behavior of most bankers.
- I have had good experience with bankers/banks so far.
- There are enough safeguards to protect clients.
- Unethical behavior by employees is usually punished by the banks.
- Bankers generally have to place the customers' interests over those of the bank.

Is there another reason, not mentioned above, why you think bankers can generally be trusted? If not, you can skip this question.

Q4. How often do you seek financial advice from your banker or financial advisor?

- I do not have a financial advisor
- Less than once a year
- Once a year
- A few times a year
- Every month or more often

Q5. Do you know whether bankers in the Netherlands/France are required to swear a professional oath of good conduct?

- I am sure that bankers do not have to swear an oath
- I am not sure, but it is unlikely that bankers have to swear an oath
- I am not sure, but it is likely that bankers have to swear an oath
- I am sure that bankers have to swear an oath

Q6. Do you think a bankers' oath should be compulsory in all European countries?

- Yes *(Go to Q7A.)*
- No *(Go to Q7B.)*

Q7A. You responded 'Yes' to the previous question. Why do you think a bankers' oath should be compulsory in all European countries?

- Because it disciplines the bankers
- Because it makes customers' interest more salient to the bankers
- For other reasons

You have responded ‘For other reasons’ to the previous question. Please explain why you think a bankers’ oath should be compulsory in all European countries.

Q7B. You responded ‘No’ to the previous question. Why do you a bankers’ oath should not be made compulsory in all European countries?

- Because people can rely on their own experience to trust bankers or not
- Because an oath does not do anything against dishonest bankers
- For other reasons

You have responded ‘For other reasons’ to the previous question. Please explain why you think a bankers’ oath should not be compulsory in all European countries.

Q8. What is your gender?

- Male
- Female
- Other

Q9. In what year were you born? (*e.g.* 1970)

Q10. What is the highest level of education you have obtained? If you are currently enrolled, please select the highest level of education you have attained.

- High school or below
- Bachelor
- Master or above

Q11. What is your current status?

- Student
- Employee
- Self-employed
- Retired
- Unemployed

Q12. What is your annual personal net income (your salary after income tax, social security contributions, and pension contributions have been deducted)?

- < €15000
- €15000 - €29999

- €30000 - €44999
- €45000 - €59999
- €60000 and more

Q13. The next question is about the following problem. In questionnaires like ours, some participants sometimes do not read the questions carefully and just click quickly on the questionnaire. This means that there are many random answers that jeopardize the results of surveys. To show that you have read our questions carefully, give ”bankers” as your answer to the following question. What is the profession indicated?

- Firefighters
- Bankers
- Doctors
- Teachers
- Police officers

End of the survey

You have completed the survey. We thank you for your time. Important: Please complete the following 3 steps to register your survey response and receive your payment.

1. Write down this survey code: *(Survey completion code here)*
2. Click on ‘Submit’ on this page to register your responses. If you do not complete this step, we will not receive your information and cannot reward you.
3. Enter the survey code into your Prolific account to register your submission.

F.2 Results of the survey and discussion

Low trust in bankers Using the survey from the Netherlands as our main case, we find low average levels of trust when it comes to the banking sector: when asked to rate their trust on a scale of 1 (‘I don’t trust at all’) to 5 (‘I trust completely’), the average rating for bankers is 2.55 (which is barely higher than in the French sample - 2.40 -, although there is no such compulsotry oath in France). Bankers rank at the bottom compared with members of parliament (who barely do better with an average of 2.61), judges (3.87), doctors (3.96), and firefighters (4.62). Our survey points to several reasons for this lack of trust: respondents believe bankers do not prioritize clients’ interest, that the banking culture is not trustworthy, and that unethical behavior is hard to detect (and therefore to punish). All summary statistics are provided in Table F1.

Low awareness of the oath Is the Dutch bankers' oath then of no impact? Our survey shows that despite an oath being compulsory, the awareness of it is limited. In fact, only 53% of our respondents were either sure or thought it is likely for such an oath exists. Whether these responses are the result of knowledge of the actual oath or simply the result of the respondents thinking that bankers "obviously" have to swear an oath given their positions is uncertain. To be able to speak to this point, we conducted a similar survey in France, where such an oath does not exist. In that survey about 28% of respondents still believe it is the case. Extrapolating this to the Dutch survey hints that perhaps the answers of a non-negligeable proportion of the 53% mentioned above do not stem from the "actual" knowledge of the oath. Responses to the other questions are much more in line across the surveys.

A correlation between oath awareness and trust in bankers The results of the survey complement our laboratory findings. While generally trust is low, there is, however, a positive link between being aware of the oath and trust in bankers. Respondents in the Netherlands who are aware of the oath hold a higher level of trust in bankers (Spearman's correlation coefficient = 0.15, $p = 0.035$). On the side of investors, our laboratory results show that being aware of the oath is not a necessary condition for trust since participants learned from their experience. This element is also reflected in our survey, as about 60% of the respondents with high trust in bankers indicated that it was because of the good experiences they have had with their bankers and banks.

In all, there are two takeaways from the field survey. First, the awareness of bankers' oath is far from being universal, even when such a practice has been legally implemented country-wide and many discussions have been conducted both in the media and at the parliament. Second, the banking profession suffers from trust issues, but they may be alleviated as a large majority of people in the survey support the implementation of a bankers' oath. Indeed, Table F1 shows that 83% of the respondents are in favor of a compulsory bankers' oath in the European Union (this is also the case for 80% of the respondents in the French sample). 60% of those who are in favor of such oath believe that it will help in making their interests more salient, and one third think that it would discipline bankers. The field experiment of Weitzel and Kirchler (2023), using reminders of the oath, shows that these beliefs are in fact well grounded.

Table F1: *Overview of responses from the Prolific survey in the Netherlands and France*

Country	Netherlands N=198	France N=197
Trust attitudes and level of awareness		
Most people can be trusted	69%	45%
Mean trust in different groups (St.Dev.)		
Doctors	3.96 (0.77)	3.96 (0.75)
Members of the parliament	2.61 (0.87)	2.22 (0.81)
Judges	3.87 (0.83)	3.37 (0.95)
Firefighters	4.62 (0.56)	4.46 (0.68)
Bankers	2.55 (0.96)	2.40 (0.95)
Three most selected reasons for ...		
Low trust in bankers	n=162	n=174
Bankers do not prioritize customers' interest	78%	84%
Banking culture is not trustworthy	60%	65%
Unethical behaviors are hard to detect	49%	40%
High trust in bankers	n=36	n=23
I have had good experiences	64%	83%
There are enough safeguards	67%	39%
Scandals are rare	53%	61%
Awareness		
I am sure that bankers do not take an oath	2.53%	20.81%
I am not sure but it is unlikely that bankers take an oath	43.94%	51.27%
I am not sure but it is likely that bankers take an oath	36.36%	22.84%
I am sure that bankers take an oath	17.17%	5.08%
Spearman's correlation between awareness and trust in bankers	0.15 ($p = 0.036$)	0.02 ($p = 0.782$)
Should bankers' oath should be compulsory in EU?		
Yes because ...	n=165	n=158
It disciplines bankers	34%	37%
It makes customers' interest salient	60%	57%
Other reasons	6%	6%
No because ...	n=33	n=39
People can rely on experience	3%	13%
Oath does not do anything	94%	79%
Other reasons	3%	8%
Socio-economic characteristics		
Male	49%	52%
Mean age (St.Dev.)	36 (9.2)	37 (10.8)
Education		
High school or below	12.63%	1.02%
Bachelor	49.49%	10.66%
Master or above	37.88%	88.32%
Status		
Student	7.58%	11.17%
Employee	76.77%	55.33%
Self-employed	8.59%	22.34%
Retired	0.51%	6.60%
Unemployed	6.57%	4.57%
Frequency of financial advice		
No financial advisor	59.60%	25.38%
Less than once a year	23.74%	50.76%
Once a year	11.62%	18.27%
A few times a year	5.05%	5.58%
Every month or more often	-	-
Annual personal net income		
< €15000	17.17%	27.92%
€15000 - €29999	22.73%	36.55%
€30000 - €44999	33.33%	21.32%
€45000 - €59999	17.68%	9.14%
€60000 or more	9.09%	5.08%

Notes: The table reports the descriptive statistics of the responses to the Prolific survey conducted with residents in the Netherlands and France. For awareness of the oath, it is coded as 1 if the respondent indicated that he or she is sure or it is likely that Dutch (or French) bankers take an oath of conduct.