I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14931

The Way People Lie in Markets: Detectable vs. Deniable Lies

Chloe Tergiman
Marie Claire Villeval

DECEMBER 2021

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14931

The Way People Lie in Markets: Detectable vs. Deniable Lies

Chloe Tergiman
Penn State University

Marie Claire Villeval
Univ Lyon, CNRS and IZA

DECEMBER 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Way People Lie in Markets: Detectable vs. Deniable Lies*

In a finitely repeated game with asymmetric information, we experimentally study how individuals adapt the nature of their lies when settings allow for reputation-building. While some lies can be detected ex post by the uninformed party, others remain deniable. We find that traditional market mechanisms such as reputation generate strong changes in the way people lie and lead to strategies in which individuals can maintain plausible deniability: people simply hide their lies better by substituting deniable lies for detectable lies. Our results highlight the limitations of reputation to root out fraud when a Deniable Lie strategy is available.

JEL Classification: C91, D01, G41, M21

Keywords: lying, deniability, reputation, financial markets, experiment

Corresponding author:
Marie Claire Villeval
Université de Lyon
GATE UMR 5824
93 Chemin des Mouilles
F-69130 Ecully
France
E-mail: villeval@gate.cnrs.fr

* We are grateful to N. Ali, G. Attanasi, I. Esponda, L. Panaccione, K. Schlag, J. Tremewan, E. Vespa, A. Wilson, S. Yuksel for very valuable feedback. We also thank participants at the Winter Business Economics Conferences in Snowbird, Utah, the Society for Experimental Finance in Salt Lake City, Utah, the ESA World Meeting in Vancouver, the North-American ESA Conference in Antigua, the Asia-Pacific ESA Conference in Abu Dhabi, the EWEBE workshop in Lyon, the LAGV Conference in Aix-en-Provence, the REBEW workshop in Reading, the meeting of the Canadian Economic Association in Banff, and at seminars at CEES, LISER, Florida State University, Middlebury College, Penn State University, the University of the Basque Country in Bilbao, the University of California at Santa Barbara and Davis, Mazarych University, and the Universities of Auckland, Bath, Goteborg, Melbourne, Michigan, Nantes (LEMNA), Paris-Dauphine, Pittsburgh, Prague, Rome (LUISS) and Vienna for useful comments. Financial support from the FELIS program of the French National Agency for Research (ANR-14-CE28-0010-01) and from IDEXLYON at Universite de Lyon (project INDEPTH) within the Programme Investissements d'Avenir (ANR-16-IDEX-0005) is gratefully acknowledged. This research has also benefited from the support of the LABEX CORTEX (ANR-11-LABX-0042) of Universite de Lyon, within the program Investissements Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR).

1 Introduction

"Capitalism is based on self-interest and self-esteem; it holds integrity and trustworthiness as cardinal virtues and makes them pay off in the marketplace, thus demanding that men survive by means of virtue, not vices." This quote from A. Greenspan (1967) highlights a central tenet of an economic system based on free trade: trust and honesty in transactions. Trust and honesty are fundamental in financial markets where investors have to rely on the ethics of banks and financial advisers that benefit from private information about the expected returns of investments (*e.g.*, Guiso et al. (2008); Gennaioli et al. (2015); Gurun et al. (2018); Zingales (2015)). In fact, such reliance is widespread: Hung et al. (2008) estimate that 73% of investors consult a financial adviser before buying shares, and Egan et al. (2019) indicate that 56% of American households ask financial professionals for advice.

Yet, financial misconduct is not rare, even when there is a fiduciary duty towards investors. In financial markets fraud takes the form of dishonest schemes, promises of unrealistic returns, book cooking or favoritism (*e.g.*, Cooper and Frank (2005); Mullainathan et al. (2012); Piskorski et al. (2015); Brown and Minor (2016); Pool et al. (2016); Anagol et al. (2017)). After building a large dataset of financial advisers in the United States from 2005 to 2015, Egan et al. (2019) found that about seven percent of advisers have misconduct records, and this percentage goes up to 15 percent at some of the largest advisory companies. Moreover, about a quarter of individuals who have misconduct records are repeat offenders. In short, fraudulent behavior is a pervasive feature of the industry.[1] This misconduct is costly: Grasshoff et al. (2017) estimate that the penalties and legal costs from misconduct cases inflicted to banks represent about USD 321 billion since 2008.[2] The cost of scandals for investors, as a result of direct or indirect investment losses, is also fairly large.[3]

Frauds are frequent not only because of their expected financial return, but also because they are detectable only to varying degrees. Some will almost surely be detected (*e.g.*, the creation of fake bank accounts, as in the 2018 Wells Fargo scandal, or Ponzi schemes like in the 2008 Madoff scam), whereas others are deniable (*e.g.*, inflated earnings announcements by companies (Roychowdhury, 2006; Beyer et al., 2010), or a purposeful increase in the complexity of financial disclosures (Zhe Jin et al., 2018)). Financial scandals have occurred in companies benefiting from a high reputation (*e.g.*, Enron, Lehman Brothers, WorldCom,

[1] For experimental evidence see also *e.g.*, Cohn et al. (2014); Gibson et al. (2017)).

[2] See also the report on misconduct risks in the financial sector by the Financial Stability Board to the G20 leaders (FinancialStabilityBoard (2017)).

[3] For example, in the Madoff scandal the direct wealth losses were estimated around $17 billion but the reduction in investment due to the trust shock was around $430 billion, implying that the losses of direct victims have represented less than 4% of the liquidation of risky assets (Gurun et al., 2018) . Graham et al. (2002) estimate that the Enron and subsequent accounting scandals led to a reduction of the U.S. GDP between 0.2 and 0.5% over a one-year period and between 1.05 and 2.5% over ten years.

Deutsche Bank).[4] This is surprising since losing reputation entails large costs. One may thus suspect that reputation affects not only the likelihood of misconduct but also the nature of lies through their degree of detectability or deniability.

In this paper we contribute to the literature on dishonesty by experimentally studying the *nature* of lies by players in the role of project managers, and evaluating how the introduction of reputation affects the kinds and frequency of lies in dynamic settings.[5] We design the "Announcement Game", a new deception game that allows players to select the type of lies they make, from detectable to deniable, and study how these choices are affected by the market institution in place. We identify a *"Deniable Lie"* strategy that maintains plausible deniability, and show that the use of this strategy responds to the presence of reputation.[6] In our setting, project managers are privately informed about the quality of their financial products, and have to attract uninformed investors to invest with them, which gives them an incentive to exaggerate the announced expected returns of investment. Investors have some chance of discovering fraudulent announcements *ex post*, which depends on the nature of the lies. Indeed, the market itself delivers *ex post* transparent information about the actual return of projects. Lies become identifiable when realized outcomes are incompatible with the announcements. In other cases, realized outcomes do not contradict the announcements, making lies *deniable*. The consequences of being identified may differ depending on whether the interaction is one-shot *vs.* repeated, as reputation may allow investors to punish detected liars but may not be so effective in discouraging deniable lies. How do people lie in markets? Does the introduction of reputation-building decrease all types of lies? Or does this market mechanism simply change *how* project managers lie, and make them shift from detectable to deniable lies?

Our game allows us to answer these questions. Precisely, the project manager receives three cards that represent a portfolio of projects. Each card has an independent 0.5 probability of displaying a star, indicating a successful project. The number of cards displaying a star is the project manager's private information. The project manager makes a cheap-talk announcement to the investor on his number of stars. After observing the announcement, the investor decides whether or not to invest his endowment with the project manager. Finally, Nature draws one of the project manager's three cards. Both players see the face of the drawn card. Whether or not there is a star on the drawn card determines the success or failure of the investment. Thus, it is mutually profitable for the project manager and

[4] Shirley (2020) further shows that many institutions are repeat offenders with multiple criminal charges.

[5] We follow Sobel (2020)'s definition of a lie as "a statement that the speaker believes is false", which differs from deception in that it does not require a model of the liar's intention or of how the audience will interpret the statement.

[6] There is a theoretical literature manipulating the detectability of lies in sender-receiver games (*e.g.,* Dziuda and Salas (2018); Balbuzanov (2019)). While in these models detectability depends on an exogenous communication technology, in our study it is endogenous and depends on the sender's strategy.

the investor if investment occurs in a good state (there are "many" stars among the three cards), but interests are not aligned when the state is bad (too "few" stars).

The message space of project managers, richer than a dichotomous "invest / do not invest" message space, allows us to identify four types of lies. **"Extreme"** lies lead to draws inconsistent with the announcement: reporting three stars when the truth is zero. These lies are detected *ex post* with certainty. **"High Risk"** and **"Low(er) Risk"** lies lead to draws that can be inconsistent with the announcement, for example reporting three stars when the truth is one or two. These lies are detected ex-post with a high (67.7%) and low(er) (33.3%) probability, respectively. **"Deniable"** lies lead to draws that are always consistent with the announcement, for example reporting two stars when the truth is zero or one. These lies can never be detected since both a blank card and a card with a star are consistent with a 2-star announcement, allowing the project manager to maintain plausible deniability. Thus, our Announcement Game allows the project manager to modulate the "intensity" of the lie he can tell. It also allows us to characterize a Deniable Lie strategy as one in which project managers announce three stars when this corresponds to the truth and, when they lie, only make deniable lies. Note that our parameters are deliberately chosen so that from the standpoint of most project managers, the monetary incentive to misreport the truth likely exceeds the intrinsic moral cost of lying. This allows us to focus on how the nature of lies (detectable *vs.* deniable) evolves depending on the market structure.[7]

We compare behavior across two main treatments. In the **"No Reputation"** treatment, at each new period each project manager is randomly rematched with an investor. In this setting, there is no possibility of reputation building with a given investor. In the **"Reputation"** treatment, pairs remain fixed throughout the session. This setting allows project managers to build a reputation (understood as fixed matching) with an investor. It allows investors to update their beliefs regarding the honesty of their manager. In equilibrium, truth-telling is not more likely with reputation than without it, but behavioral conjectures predict that project managers may react to the threat of punishment by investors.

Our results show evidence of widespread dishonesty: over 92% of subjects lie at least once. Regarding the nature of lies, absent reputation, up to 97% of subjects who lie make lies that can lead to detection. However, the introduction of reputation leads to a major change in the nature of lies: detectable lies become infrequent, and project managers shift towards a Deniable Lie strategy, which maintains plausible deniability, so as to not be detected as liars by the investors with whom they are in fixed relationships. Indeed, when in fixed matches, investors frequently punish project managers who are caught lying by not

[7]With this choice we believe that we parallel the world outside the laboratory, where the gains from misconduct may be considerably larger than the intrinsic cost of engaging in such behavior. Further, as Egan et al. (2019) show, individuals who get caught, are often simply re-hired by other firms, meaning that the negative consequences of being caught may not be as high as one may think.

investing with them in future periods, and we find that the reputation mechanism practically eliminates Extreme and High Risk lies, and reduces the relative frequency of Low(er) Risk lies. However, it has no discernible effect on the frequency of Deniable lies that can only be detected probabilistically after a large number of interactions under various states of Nature, and are therefore much harder to punish. Thus, reputation does not necessarily make project managers more honest in all situations, but instead leads them to change the nature of their lies so as to lower their likelihood of being identified as liars. Such a Deniable Lie strategy is, in contrast, seldom used in the No Reputation treatment.

After showing that project managers' earnings depend on their lying behavior, we examine other possible reasons why project managers. Although exaggerated announcements could in theory help risk averse individuals invest in situations where they should, the analysis of strategies and earnings only partially supports the paternalistic view according to which players lie to the benefit of investors (this may apply to no more than 15% of liars).[8] We also reject that lying behavior is driven by efficiency concerns or a high score in manipulativeness. Instead, lying behavior is dictated by the institution in place. As a result, in all treatments investor earnings are far below what they would be if markets were transparent. Nevertheless, we show with a third treatment, the **"No Communication"** treatment, that despite the high frequency of lying, communication yields higher returns for investors than a market in which project managers cannot send any messages. This is because in some cases signals are informative.

Our contribution is threefold. First, we propose a novel sender-receiver game in which the informed party can endogenously choose how detectable their lies will be, ranging from lies that will be detected *ex post* with certainty, to lies that are deniable. Second, we show that in the presence of reputation, project managers shift to a Deniable Lie strategy. Third we find that with reputation, while investors punish those that are caught in a lie, a project manager who uses deniable lies is not punished. In fact using the Deniable Lie strategy increases payoffs. These findings have crucial implications for increasing trust-building in financial companies. They highlight the importance of not only reinforcing reputation mechanisms, but also improving the truthfulness of communication, notably through the auditing procedures implemented by regulatory agencies to protect consumers against lies that are deniable. A complementary line of action to limit the prevalence of deniable lies might be to increase the moral cost of such lies by increasing the personal responsibility of advisers.

[8] In a separate context, Ambuehl et al. (2021) show that paternalistic motivations can explain behavior. However, paternalism there is defined as imposing one's own preferences on others, while here its definition does not require the mapping of one's own preferences: it is having others' best financial interests at heart.

2 Contributions to the Literature

Our paper falls within the literature on how individuals report private signals in asymmetric information settings. We are the first to propose a deception game with partial *ex post* information that allows players to endogenously select the type of lies they make, from some that have varying degrees of detectability to others that are deniable, and study how these choices are affected by reputation-building in a dynamic setting.

We contribute to the large literature on the reporting of private signals in asymmetric information games, a literature that spans both individual decision-making choices (cheating tasks) and games in which players interact (deception games). In contrast to our study, the recent focus of this literature has been on the identification of the moral costs of lying, primarily using static games without detectability. The main results of cheating games, inspired by the die-rolling paradigm, are that even when there is no scrutiny, no negative externalities to lying and lies are profitable, not all individuals lie, most liars do not lie in full, and the size of lies does not increase with incentives (see the meta-analysis of Abeler et al. (2019) and the references therein). Likewise, in deception games with strategic information transmission *à la* Crawford and Sobel (1982), although lying is more substantial than in cheating games and is an equilibrium prediction, subjects often tell the truth.[9] In contrast with cheating games, however, raising the level of incentives increases lying (Gneezy, 2005; Sutter, 2009; Erat and Gneezy, 2012). The main models put forward to explain these patterns are that (1) a fraction of people suffer from lying aversion because of an intrinsic preference for truth-telling;[10] and (2) people care about the reputational cost of lying.[11]

Our Announcement Game differs from standard deception games in several respects. The payoff structure of the game is common knowledge and players can manipulate the deniability of lies, which allows only some lies to be detected *ex post*. This allows us to focus on lying strategies characterized by different levels of *ex ante* and *ex post* detectability and deniability. By allowing for a wider breadth of lies, we are able to identify a Deniable Lie strategy and show how it reacts to the market environment individuals face. Moreover, although we share with this literature an interest for the reputational cost of being identified as a cheater, our aim is not to identify the psychological cost of suspicion, but how its economic cost changes the nature of lies, depending on whether reputation building is possible or not. Because our design provides relatively high monetary incentives for lying,

[9] *E.g.*, Austen-Smith (1993); Blume et al. (1998); Krishna and Morgan (2001); Battaglini (2002); Gneezy (2005); Cai and Wang (2006); McGee and Yang (2013); Vespa and Wilson (2016); Rantakari (2016); Li et al. (2016); Schmidbauer (2017); Ederer and Fehr (2017). This does not imply that truth-telling is the norm, see Wilson and Vespa (2020). For a recent survey on strategic transmission, see Blume et al. (2020).

[10] *E.g.*, Ellingsen and Johanesson (2004); Sánchez-Pagés and Vorsatz (2007); Vanberg (2008); Hurkens and Kartik (2009); Shalvi et al. (2011); Fischbacher and Föllmi-Heusi (2013).

[11] On perceived cheating aversion see Kajackaite and Gneezy (2017); Dufwenberg and Dufwenberg (2018); Gneezy et al. (2018); Abeler et al. (2019); Khalmetski and Sliwka (2019).

we anticipate that the way market mechanisms affect extrinsic motivation constitutes the primary force behind the decision to change the nature of one's lies.[12]

Second, our work is connected to the literature on communication with evidence. This literature has its roots in the theoretical works of Grossman (1981) and Milgrom (1981), but the experimental literature on disclosure games is relatively small (e.g., Forsythe et al., 1999; Dickhaut et al., 2003; Benndorf et al., 2015; Sánchez-Pagés and Vorsatz, 2009; Hagenbach and Perez-Richet, 2018; Li and Schipper, 2020; Jin et al., 2021). For example, in Hagenbach and Perez-Richet (2018) senders can fully disclose their private information to uninformed receivers, partially disclose it, or withhold information, which allows them to deceive the receivers if they are envious. In Jin et al. (2021), senders also choose whether or not to disclose private information; receivers are found to be not sufficiently skeptical about undisclosed information. In contrast with these games that allow for deception but not lying, if our project managers want to hide their private information they have to misrepresent it.[13] Moreover, in our game the information that is disclosed is not verifiable and only partial information is shown and after the investor takes action.

Our Announcement Game also allows us to bring important nuances to the study of how market mechanisms support honesty when information is asymmetric. The literature on credence goods has shown that among reputation, verifiability, liability, competition and the interaction thereamong, only liability (the obligation for the seller to provide sufficient quality) leads to significantly more honesty.[14] Considering instead experience goods and a trust game, Huck et al. (2012) found that as long as trustors can identify trustees, competition among trustees is sufficient to achieve efficiency; adding reputation through the full history of play offers no improvement. Lacking from these studies is the ability of the informed party to modulate how they lie. We show that with a richer set of lie types, as is even more the case outside the laboratory, reputation fails to root out fraud: liars adapt the nature of their lies. While it is reassuring that reputation eliminates frauds for which the informed party is very likely to be exposed, fraudulent announcements persist when they are deniable. By nature, these lies are those that are difficult to identify by uninformed parties (and also by empirical studies), and lead to large negative consequences for them.

[12]We show that "how much to lie" depends on the costly consequences to getting caught, more than on the distance between a reported outcome and the truth. The literature has discussed the size of lies in the absence of punishment, showing that partial lying is more frequent when the experimenter knows the truth. Here, equal sizes of lies have different probabilities of detection and we show that deniable lies are uniform across treatments while detectable lies are not, even if the size of the lie in its outcome dimension is identical.

[13]It has been found that vague messages increase lying (Serra-Garcia et al., 2011) because lying is less costly (see Turmunkh et al. (2019) with *ex post* verifiability). Also, introducing punishment does not increase truth-telling when messages are evasive -pretending not to know- (Khalmetski et al., 2017) (see also Sanchez-Pages and Vorsatz (2009)). Here, announcements cannot be concealed and messages are precise.

[14]See Emons (1997); Charness and Garoupa (2000); Bohnet et al. (2005); Dulleck et al. (2011); Balafoutas et al. (2013); Beck et al. (2014); Mimra et al. (2016); Rasch and Waibel (2018); Feltovich (2019).

3 Design, Procedures and Predictions

3.1 Design

3.1.1 The Announcement Game

In order to study the way people lie in markets, we propose a new game, the Announcement Game, in which lying is an equilibrium outcome in a "standard" model. The Announcement Game has a finite, but unknown, horizon. Subjects are informed that they will play this game for a minimum of 20 and a maximum of 40 periods. We introduced such uncertainty to avoid end game effects.[15] The actual number of periods was actually fixed at 27 in all sessions. It is common knowledge that the total number of periods and the period randomly drawn for payment have been determined before the beginning of the experiment. At the end of the session subjects learn which period was randomly selected for payment; the decisions made in this period determine the subject's payoff in the game. The game was played in two between-subject treatments: the No Reputation treatment (which constitutes our baseline) and the Reputation treatment. We describe each treatment successively.

No Reputation Treatment: Half of the subjects are randomly assigned the role of a project manager ("participant A" in the instructions, see Appendix A) and the other half the role of an investor ("participant B"). Roles are fixed. In each period, a project manager is matched with an investor and pairs are randomly rematched at the beginning of each new period. We now describe the timing of each period.

(i) First, Nature randomly draws a set of three cards for each project manager. The cards convey information on the quality of the manager's projects. Each card has a (independent) 0.5 probability of displaying a star, which indicates a successful project. If the card does not display a star, it is blank. Only the project manager can observe his three cards and see how many display a star. The total number of stars τ is then in the $\{0, 1, 2, 3\}$ set.

(ii) Then, the project manager sends a cheap-talk message m to the investor regarding the number of cards that display a star, where $m \in \{0, 1, 2, 3\}$. Although this possibility is not stated in the instructions, m can differ from τ: a project manager can misreport the number of stars observed to motivate the investor to invest.

(iii) The investor receives an endowment of 100 tokens. After observing m, he decides on an action $a \in \{Invest, Not\ Invest\}$. If $a = Invest$, the entire endowment is invested.

(iv) Nature randomly draws one of the project manager's three cards, which determines the success of the project, $\theta(\tau) \in \{Star, No\ Star\}$. If the selected card displays a star, the project is a success; if the card is blank, it fails.

[15] We acknowledge that such uncertainty may introduce an additional source of heterogeneity across subjects, but this is kept constant across treatments and we are mainly interested in treatment comparisons.

(v) Finally, both the project manager and the investor learn θ, regardless of whether the investor invested or not, and receive payoffs. A history box is also displayed on the subjects' screen. This history box lists (m, a, θ, π^i) for each past period. The past values of τ are only part of the history box for the project managers and investors never learn τ.

Figure 1 displays the timeline of the game.

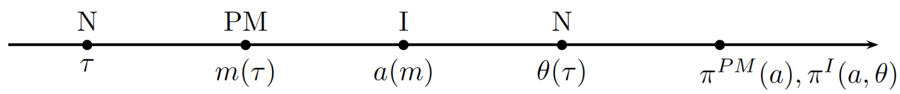


Figure 1: Timeline of the Game

The project manager's payoff is state independent and determined entirely by a:

$$\pi^{PM}(a) = \begin{cases} 30 & \text{if} \quad a = \text{Not Invest} \\ 230 & \text{if} \quad a = \text{Invest} \end{cases}$$

The project manager earns a fixed amount of 30 tokens plus 200 tokens if the investor invests, regardless of θ, the quality of the selected project.[16]

The investor's payoff is state dependent and also depends on a:

$$\pi^{I}(a, \theta) = \begin{cases} 100 & \text{if} \quad a = \text{Not Invest} \\ 30 & \text{if} \quad (a, \theta) = (\text{Invest, No Star}) \\ 300 & \text{if} \quad (a, \theta) = (\text{Invest, Star}) \end{cases}$$

If he does not invest, the investor earns his initial endowment (100 tokens). If he invests in a failed project, his net payoff is 30 tokens. If he invests in a successful project, his net payoff is 300 tokens, triple of the amount invested. Thus, interests are aligned when the initial state of nature is good (*i.e.*, so long as $\tau \geq 1$ since a perfectly informed risk neutral investor should rationally invest if the set of cards includes at least one star). But they are unaligned when it is bad: when τ=0, project managers would like the investor to invest, but by doing so the investor would lose most of his endowment.[17]

[16] This captures both the high share of variable pay in the earnings of advisers in financial institutions and the fact that variable pay depends on the ability to sell given products, not on the success of these products.

[17] These parameters are such that only very risk averse investors should not invest, while moderately risk averse, risk neutral or risk loving individuals should invest regardless of the announcement. This feature limits the potential benefit of lies. We justify it by two reasons. First, this feature allows us to observe a full typology of lies in a parsimonious design. Suppose investing was only profitable in expectation if the true number of stars were at least 2. Then, a moderately risk-averse subject would only invest conditional on a 3-star (truthful) announcement. This would lead subjects to make detectable lies to persuade investors

Reputation Treatment: This treatment allows us to isolate the impact of reputation-building on lying behavior and investment. Rules are similar to those of the baseline except that the investor and the project manager remain in a fixed match throughout the Game.

3.1.2 The Truthful Announcement Game

The outcome of the investment involves compound lotteries and we know that individuals may face difficulties calculating their return (*e.g.*, Abdellaoui et al. (2015)). To facilitate subjects' understanding and to better understand their risk preferences, before subjects play the Announcement Game we implemented a simplified version of it in which subjects are paired with a truth-telling computer.[18] All subjects play the role of an investor and they play 16 such periods as practice.[19]

After the practice periods, subjects make decisions that can each matter for payment. They have to decide on whether to invest or not in *each* of four scenarios corresponding to $\tau = (0, 1, 2, 3)$. To determine payoffs in the Truthful Announcement Game, at the end of the session the program randomly selects one of these scenarios to count for payment.

3.1.3 Social Preferences and Other Individual Characteristics

In the Announcement Game, decisions may be influenced by social preferences. To assess whether the distribution of social preferences is similar across treatments, subjects play an Allocation Game directly following the Announcement Game. In each round, subjects are paired randomly. In each pair one subject has the role of the X player and the other has the role of the Y player. Roles and partners are assigned randomly and independently from the roles and pairs assigned in the Announcement Game. All the subjects make 15 decisions as player X under the veil of ignorance. Then, a random draw assigns a role to each subject and selects one decision for payment. Only the decisions of the X players matter for payment. In all cases, the order of decisions is randomized at the individual level. In each round, the

to invest, and deniable lies would serve no purpose. Restoring the attractiveness of deniable lies would then suppose to increase the total number of cards, making the game overall more complex. Second, this feature limits the harm a lie makes on investors, which increases the likelihood of observing misconduct in the Baseline, leaving more room for reputation-building to affect behavior.

[18] It is simplified in the sense that in this treatment the lottery is no longer a compound one, and subjects do not have to form beliefs on how truthful someone is.

[19] The rules are the same regardless of the treatment implemented in part 2, and subjects are not aware of the rules for part 2 when they play the Truthful Announcement Game. They are informed on the probability of observing each number of stars among the three cards. Moreover, to facilitate learning, all the subjects can experience the distribution of probabilities in the practice periods: in two periods the three cards show no stars, in six periods they show one star, in six periods they show two stars, and in two periods they show three stars, in random order.

X player has to choose between two allocations that determine payoffs for himself as well as the person he is matched with. Details are provided in Table B1 in Appendix D.

Finally, we administer the Machiavellianism (Mach-IV) test (Christie and Geis, 1970) to collect measures of manipulativeness (see Appendix A). Then, subjects receive a feedback on their payoffs in each part and answer to a standard socio-demographic questionnaire.

3.2 Procedures

The experiment was run at the GATE-Lab in Lyon, France. We ran 8 sessions for the main experiment: 4 sessions with 84 subjects in the No Reputation treatment, 4 sessions with 78 subjects in the Reputation treatment.[20] All 234 participants (58.55% of female) were recruited via HRoot (Bock et al., 2014), mainly among students from local engineering, business and medical schools. The experiment was programmed using z-Tree (Fischbacher, 2007).

Upon arrival, subjects randomly drew a ticket from a bag indicating their terminal number. The instructions for each part were distributed at the beginning of the relevant part and read aloud. In Part 1, subjects played the Truthful Announcement Game. In Part 2 they played 27 periods of the Announcement Game. In Part 3 they participated in the Allocation Game. In Part 4 they took the Mach IV test. Finally, they answered a standard socio-demographic questionnaire. Sessions lasted about 1.5 hour.

Subjects were paid the sum of their payoffs in the first three parts: the payoff for the payoff-relevant decision in the Truthful Announcement Game, the payoff for one randomly chosen round in the Announcement Game, and the payoff for one randomly chosen round in the Allocation Game. Feedback on payoffs in each part was given only at the end of the session. The average earnings were 17.9 Euros, including a 5-Euro show-up fee. Earnings were paid in cash and in private in a separate payment room.

3.3 Predictions

As in other cheap talk games, there are multiple equilibria in our Announcement Game. With risk-neutral players with selfish preferences and no behavioral types, truth-telling cannot be supported. Indeed, project managers do not make announcements that lead to no investment, and a risk neutral investor should invest in all periods.[21] Note that if we

[20] We also ran a robustness treatment with 4 sessions and 72 subjects, which we describe in Appendix F.2.

[21] In the last period, project managers will not send a signal that reduces the likelihood of an investment, and telling the truth (which could mean sending a message m of 0) leads to a lower payoff than lying. The same rationale applies to the periods preceding the last one. In expectation, without receiving any credible information from the project manager, investors are always better off investing than not. Indeed, it is profitable to invest when the project manager's portfolio contains at least one star since the expected profit from investing with τ=1 is 120 (=(300*0.33+30*0.67)), which is higher than the certain payoff of 100

assume that subjects have preferences for efficiency, then investment also occurs in every period. Indeed, investment coincides with the efficient choice, regardless of the true number of stars. Thus, preferences for efficiency also lead to investment in every period and the project manager announcements are irrelevant (and can thus be truthful in equilibrium).

However, we may observe more honesty from project managers than would be predicted by standard preferences. Indeed, behaviorally, investors may react adversely to the realization that they have been lied to, making their behavior history-dependent, and leading them to seek to punish detected liars. Project managers may anticipate punishment from the investors they have hurt and who may stop investing with them.[22] In this case they may adapt their behavior to the environment they face, leading to differences in how they lie when Reputation is present. We also note that, while we deliberately chose parameters so that the gains from lying for the project manager are high (earning 230 instead of 30) relative to the potential cost for the investor (since investing is profitable in expectation regardless of the announcement), project managers who have a very strong preference for honesty or perceived cheating aversion (e.g., Dufwenberg and Dufwenberg, 2018; Gneezy et al., 2018; Abeler et al., 2019) may also more truthfully report their cards.

We highlight one particular strategy that maintains plausible deniability, which we term the Deniable Lie strategy. With this strategy, project managers are truthful when they face 3 stars, and otherwise when they lie only make lies that can be denied. The Deniable Lie strategy is part of a partially informative equilibrium in the Reputation treatment, since with this strategy the investor invests in all periods and there is no incentive for the project manager to change strategy.[23] In this strategy, project managers never make announcements that can be incompatible with Nature's draw. We now introduce our first conjecture.

Behavioral Conjecture for Project Managers: In the Reputation treatment, project managers primarily lie in a way that maintains plausible deniability. There are far fewer detectable lies in the Reputation treatment than in the No Reputation one. In the No Reputation treatment, the Deniable Lie strategy is less frequent and any kind of lie may occur.

if one does not invest. The expected number of stars being 1.5, investors should always invest.

Note that in our settings, babbling equilibria are Pareto-efficient (but our typology of lies does not depend on this property). This is in contrast with canonical models such as Crawford and Sobel (1982) where babbling equilibria are inefficient for both sender and receiver.

[22] This conjecture builds on the abundant literature on trust and public goods games that showed how fixed matching increases the punishment threat and the level of cooperation in anticipation of such punishment. See, e.g., Brandts and Charness (2003); Sánchez-Pagés and Vorsatz (2007); Wilson and Wu (2017); Bernard et al. (2018); for surveys, Thöni (2017); Villeval (2021).

[23] The equilibrium in which project managers use the Deniable Lie strategy is one in which risk neutral and moderately risk averse investors invest in every period; this equilibrium is payoff-equivalent to a babbling equilibrium even though, as opposed to babbling equilibria, some information is transmitted here.

While both the "standard" theory or preferences for efficiency predict that any given investor should always invest (assuming only moderate risk aversion) or never invest (if their level of risk aversion is high), we expect a drop in investment in the period following a detected lie in the Reputation treatment. There should be less of a period-to-period difference in the No Reputation treatment where directed punishment is not possible.

Behavioral Conjecture for Investors: In the Reputation treatment, investors punish project managers caught lying by not investing with them in the next period.

4 Results

We report the results of non-parametric tests for which we average values at the individual level; tests are run using a single observation per individual and they are two-sided. Unless otherwise stated, we use Mann-Whitney rank-sum tests for comparisons across treatments and Wilcoxon matched-pairs signed-ranks tests or tests of proportions for comparisons within treatments. We confirm our results via the use of regressions with standard errors clustered at the session or individual level. We begin by analyzing project managers' behavior through a typology of lies. Then, we examine investors' punishment behavior. We next explore the project managers' monetary and non-monetary motivations for lying. We next shift to the investors' decision errors and earnings.

4.1 Project Managers' Communication Strategies

4.1.1 Frequency of Misreporting

At the individual level, very few subjects always lie when they get fewer than three stars (4.8% in the No Reputation treatment and none in the Reputation treatment, $p = 0.168$). At the aggregate level, a decent share of subjects start the Announcement Game by telling the truth on what they observed: only 21.4% of the project managers lie in the first period in the No Reputation treatment and 15.3% in the Reputation treatment ($p = 0.484$).[24] However, if full lying is extremely rare and if telling the truth happens initially, the benefits of lying quickly exceed the moral cost of misbehavior, as expected given our parameter choices. Indeed, 73.8% of the project managers in the No Reputation treatment and 69.2% in the Reputation treatment ($p = 0.648$) lie at least once in the first five periods, and the vast majority of project managers (83.3% and 97.4%, respectively; $p = 0.034$) lie by the end

[24] Restricting to subjects who saw fewer than 2 stars in the first period does not change this. In this case the percentages are 31.8% and 19.0% (p=0.337).

of the game (meaning that only 16.7% and 2.6% never lied, respectively).[25]

Overall, the fraction of untrue announcements is 50.4% in the baseline and 30.5% when reputation is introduced ($p < 0.001$), which is relatively high since the denominator includes cases in which project managers observe two or three stars. The empirical and expected distributions of announcements are displayed in Table 1. In both treatments, the fraction of "low" announcements (0 or 1 stars) is significantly lower than the fraction of "high" announcements (2 or 3 stars). They should be equal if subjects told the truth ($p < 0.001$).[26]

Table 1: Distribution of Announcements by Treatment.

Treatments	No Reputation	Reputation	*Expected Distribution under truth-telling*
0 Stars Announced	3.7%	6.7%	*12.5%*
1 Star Announced	11.4%	21.3%	*37.5%*
2 Stars Announced	47.7%	53.5%	*37.5%*
3 Stars Announced	37.2%	18.5%	*12.5%*
Total	*100%*	*100%*	*100%*
Mean Announcement	2.2	1.8	*1.5*

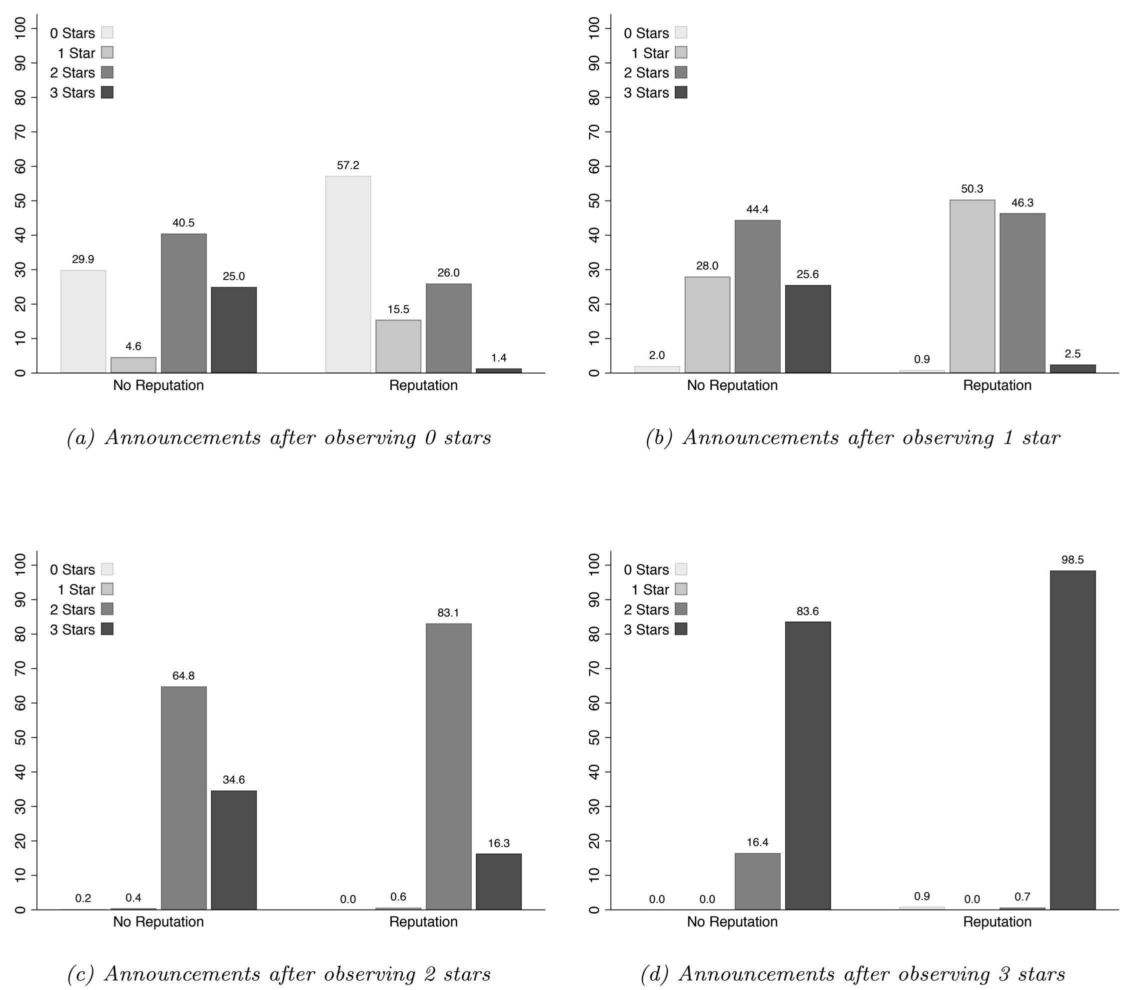
The *intrinsic* motivation for being honest and seen as honest should be constant across treatments, while strategic considerations may lead to treatment differences. To evaluate how reputation impacts announcements, we look at which announcements project managers choose in each treatment when faced with different numbers of stars. This is shown in Figure 2 where we display the distribution of announcements after observing 0, 1, 2 and 3 stars.

The data show strong evidence that lies vary based on both the treatment and the true number of stars observed. Indeed, in the Reputation treatment, lies of magnitude 1 are more frequent when a project manager observes 1 star than when they observe 2 stars (46.3% *versus* 16.3%, $p < 0.001$). The same is true of lies of magnitude 2 after observing 0 and 1 stars (26.0% *versus* 2.5%, $p = 0.001$) Similarly, in the No Reputation treatment, lies of magnitude 1 are significantly less likely after observing 0 stars than they are after observing 1 or 2 stars (4.6% *versus* 44.4% and 34.6%, $p < 0.001$ in both comparisons). Looking across treatments, lies of magnitude 2 when a project manager sees 1 star are far more common in the No Reputation treatment than in the Reputation one (25.6% *versus* 2.5%, $p < 0.001$). Differences between treatments also appear when considering the percentages of subjects who always lie after observing fewer than two stars (38.1% in the No Reputation treatment but only 2.6% in the Reputation treatment, $p < 0.001$), or after

[25] Figure C1 in Appendix C shows the relative frequency of lies over time. This frequency becomes stable very quickly.

[26] Here, we calculate the probability of an announcement greater or equal to 2 for each subject and then compare the sample proportion to 50% using a test of proportion.

observing 0 stars (63.2% in the No Reputation treatment and 21.6% in the Reputation treatment, $p < 0.001$). Figure 2 also shows that not all announcements are fraudulent, even in the absence of reputation, and even when the true number of stars is low. For example, in the No Reputation treatment 29.8% and 28.1% of the announcements following the observation of 0 and 1 stars, respectively, are truthful. Those numbers are 57.2% and 50.3% when reputation is present (and statistically different than in the absence of reputation; $p = 0.001$ and $p < 0.001$).



(a) Announcements after observing 0 stars

(b) Announcements after observing 1 star

(c) Announcements after observing 2 stars

(d) Announcements after observing 3 stars

Figure 2: Distribution of Announcements Conditional on the Observed Number of Stars, by Treatment.

4.1.2 A Typology of Lies

Moving towards better understanding project manager strategies, we shift our focus to those subjects who lied at least once over the course of the game. We categorize lies into four mutually exclusive types and analyze their frequency. **"Extreme"** lies are lies that will be detected *ex post* with 100% certainty (*i.e.*, announcing 3 stars when one has none); **"High Risk"** lies are lies that will be detected *ex-post* with probability 2/3 (*i.e.*, announcing 3 stars when one has only 1); **"Low(er) Risk"** lies are lies that will be detected *ex-post* with probability 1/3 (*i.e.*, announcing 3 stars when one has only 2); **"Deniable"** lies are lies that cannot be detected *ex-post*, *i.e.*, announcing 2 stars when one has observed fewer than 2, or announcing 1 when one has observed none.[27]

Table 2 displays the fraction of subjects who engage in each type of lies at least once and the fraction who make detectable lies (Extreme, High Risk and Low(er) Risk) at least once.[28]

Table 2: A Typology of Lies – Fraction of Subjects who Engage in Each Type of Lie at Least Once.

Treatments	No Reputation	Reputation
Extreme Lies	44.1%	2.8%
High Risk Lies	62.9%	13.2%
Low(er) Risk Lies	77.1%	50.0%
Deniable Lies	85.7%	97.4%
Detectable Lies	82.9%	50.0%

Notes: The Table displays the fraction of subjects who engage in each type of lies at least once. For example, in the No Reputation treatment 44.1% of the project managers make at least one Extreme lie over the course of the game. The total of each type of lie is not equal to 100% because each project manager may use several categories of lies at least once.

Table 2 reveals large cross-treatment differences. Absent a reputation mechanism, 82.9% of project managers who lie at least once make detectable lies, that is they announce 3 stars when they in fact face fewer than 3 stars, whether in the form of Extreme, High Risk or Low(er) Risk lies. This means that they take actions that can result in being detected as a liar by the investor they are matched with. In contrast and consistent with our Behavioral Conjecture for project managers, in the Reputation treatment a large fraction of subjects adapt their lying so as to not be identified as a liar. Among the subjects who make at least one lie, the fraction of those who make at least one detectable lie is significantly lower in the Reputation treatment than in the No Reputation treatment (50% *vs.* 82.9%; $p < 0.001$ with a test of proportions), and these lies are primarily Low(er) risk lies.

Reputation leads to a dramatic decrease in lies, but this is not uniform across all types of lies. Indeed, also in line with our behavioral conjecture, the impact of reputation is

[27] There is another category of lies: "Downward" lies, *i.e.*, announcing fewer stars than actually observed. For the sake of concision we do not comment on them, as they seldom happen (see Figure 2).

[28] Table D1 in Appendix D presents these statistics without restricting them to subjects who lie at least once. Unsurprisingly, given that the vast majority do lie, there are no substantive differences between the two tables.

dramatic in terms of Extreme lies, as it practically eliminates them (only 2.8% of the project managers make at least one extreme lie in the Reputation treatment, compared to 44.1% in the No Reputation treatment, $p < 0.001$). While a non-negligible fraction of project managers engage in High Risk lies at least once when they can build a reputation, the impact of reputation remains dramatic (13.2% *vs.* 62.9% in the Reputation and No Reputation treatments, respectively; $p < 0.001$). These patterns continue for Low(er) Risk lies, though such lies are no longer rare even when project managers can build a reputation (50% *vs.* 77.1% in the Reputation and No Reputation treatments, respectively; $p = 0.016$). Importantly, Table 2 shows that in both treatments a large majority of subjects engage in Deniable lies at least once (97.4% *vs.* 85.7% in the Reputation and No Reputation treatments, respectively; $p = 0.070$).

As a complement of the previous analysis, Table 3 reports the fraction of times subjects who lie at least once make Extreme, High Risk, Low(er) Risk and Deniable lies in each treatment, appropriately conditioned on the true observed number of stars (for example, for Extreme lies we look at cases when a subject actually saw no stars and calculate the fraction of times that subject made an extreme lie).[29]

Table 3: A Typology of Lies – Relative Frequency of Each Type of Lie.

Treatments	No Reputation	Reputation
Extreme Lies	27.9%	1.4%
High Risk Lies	30.7%	2.5%
Low(er) Risk Lies	41.5%	16.7%
Deniable Lies	52.1%	46.7%

Notes: These data focus on subjects who lie at least once. The Table displays the fraction of times subjects make each type of lies in each treatment, conditioned on the true observed number of stars. For example, in the No Reputation treatment project managers make Extreme lies 27.9% of the times when they observed 0 stars. The total of each column is not equal to 100% because the prevalence of each category of lie is conditioned on the true observed number of stars. For ex., Extreme lies only refer to cases in which 0 stars were observed, while Low(er) Risk lies refer to cases in which 0, 1 or 2 stars were observed.

The analysis of the prevalence of lies is consistent with the observations relative to the extensive margin. It confirms the dramatic impact of reputation on the likelihood of making detectable lies: in the Reputation treatment, the prevalence of Extreme lies is extremely low (1.4%), while they occur 27.9% of the time in the No Reputation treatment. Similarly, the prevalence of High Risk lies in the Reputation treatment is 2.5% compared to 30.7% in the No Reputation treatment, and the respective percentages for the Low(er) Risk lies are 16.7% and 41.5% (the highest p-value in all pairwise comparisons: 0.002). When the risk of detection decreases, the frequency of lies increases and when there is no more risk

[29]See Table D1 for equivalent statistics without restricting the sample to subjects who lie at least once.

of immediate detection, the prevalence of lies is very high in both treatments (46.7% and 52.1% in the Reputation and the No Reputation treatments, respectively; $p = 0.258$).[30]

Finally, the Deniable Lie strategy (be truthful when the truth is 3 stars, otherwise only lie in a deniable way) is much more common in the Reputation treatment (where it is employed by 47.4% of project managers) than in the No Reputation treatment (where only 8.6% of subjects employ this strategy). This suggests that reputation does not make project managers more honest in a fundamental way, but instead leads them to play strategically with the deniability of lies.

Supporting our behavioral conjecture, Result 1 summarizes the analysis of project managers' announcements.

Result 1: (1) In the Reputation treatment, 47.4% of the project managers employ the Deniable Lie strategy, in which they only make lies that for sure cannot be detected, and truthfully announce 3 when they face 3 stars. The fraction of project managers who employ that strategy in the No Reputation treatment is significantly smaller at 8.6%. (2) Very few Extreme and High Risk lies occur in the Reputation treatment. These fractions are significantly lower than in the No Reputation treatment, and this is also the case for Low(er) Risk lies. (3) Consequently, reputation reduces the fraction of project managers who engage in detectable lies, in particular the most egregious ones. However, it does not reduce the frequency of Deniable lies, and, at the intensive margin increases them.

4.2 Punishment of Detected Lies

We now show that project managers who are caught lying face consequences in the form of future non investment. In the No Reputation treatment investors are frequently aware that they have been lied to. While fewer liars get caught with reputation than without it, the fraction of liars who get caught is substantial at 34.2%, compared to 80.0% in the No Reputation treatment ($p < 0.001$; test of proportions). Furthermore, in addition to lowering the instances of detected lies, reputation delays detection, though this detection is still relatively early in the game: on average, periods 11.9 in the Reputation treatment *vs.* 6.8 in the No Reputation treatment ($p < 0.012$; Fisher exact test).

How are detected project managers punished in future periods? Figure 3 shows investment rates in a period by distinguishing the three cases that may have happened in the

[30] These analyses are confirmed by the regressions reported in Appendix Table E2. Moreover, Appendix Tables D2 and D3 in Appendix D mirror Tables 2 and 3 but are conducted on each block of nine periods. These tables show that the proportion of project managers who engage in the different types of lies over the three blocks is relatively stable in both treatments. However, in the Reputation treatment we note an increase in the intensity with which they make each of those lies. For example, the relative frequency of Deniable lies is 35.9% in the first block, and 55.2% in the third. There is less movement in the No Reputation treatment.

previous period: no star was drawn by Nature and a lie was detected, no star was drawn but no lie was detected, a star was drawn.

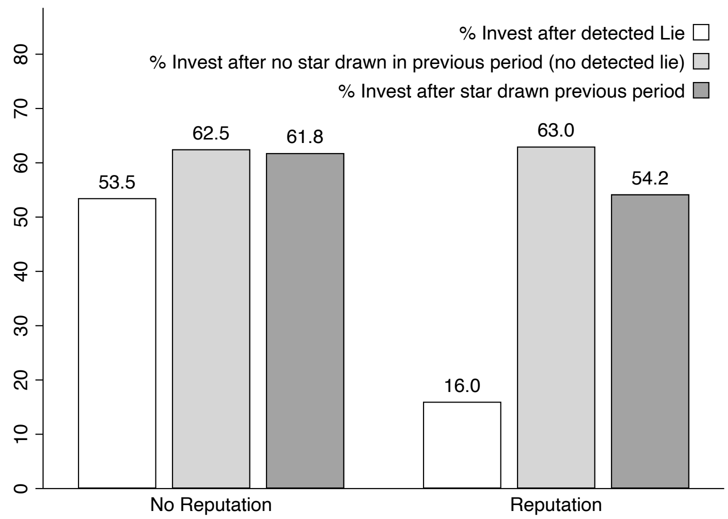


Figure 3: Investment Rates in Periods Directly Following a Detected Lie, a Star drawn, or No Star Drawn.

In the No Reputation treatment, following a detected lie the average investment rate is 53.5%, compared with a 62.5% investment rate in periods after the final draw showed no stars but no lie was detected; the difference is not significant ($p = 0.120$). Thus, investors do not punish the current project manager for a previous bad outcome if the project manager in the previous match was caught lying, showing no evidence of immediate indirect punishment. Further, when no lie was detected, the investment rates are similar whether the final draw showed a star or not (61.8% and 62.5%), indicating no reaction to the payoff outcome itself in this treatment ($p = 0.795$).

In contrast, when matching is fixed, following a detected lie the average investment rate is 16.0% compared with 63.0% in periods in which the final outcome was also bad but no lie was detected ($p = 0.011$). This means that for a given bad outcome, lie detection leads to a lower likelihood of investing in the next period. This is the case even though the average announcement following a detected lie is no lower than the average announcement following an announcement not detected as a lie ($p = 0.861$). This provides an explanation as to why project managers in the Reputation treatment only seldom make Extreme and High Risk lies: investors punish liars, although in expectation it is very costly for them to do so.

Finally, we highlight an important aspect of the cross-treatment comparison: investors invest far less with a project manager following a detected lie when they are in a fixed match with them (16.0%) compared to when matching is random (53.5%) ($p < 0.001$).

Punishment can be more than a simple one-time decision to not invest. Indeed, another form of punishment of lies is “permanent” exit, which in the No Reputation treatment represents a “punishment of the profession”, as opposed to the long-term punishment of a specific manager as in the Reputation treatment. For each subject, we identify the final period they invested in. The fraction of permanent exits is higher when reputation is present: 7.7% of the investors exit before the 20th period and 23.1% before the 25th period, against 4.8% and 14.3% in the No Reputation treatment.

To explore further the dynamics of punishment over time, Table 4 shows the coefficients from random-effects Probit models in which the dependent variable is the decision to invest after seeing a 2-star announcement (left panel) or a 3-star announcement (right panel). In the No Reputation treatment, the independent variables include a time trend, the running number of detected lies (a 3-star announcement followed by the draw of a blank card), the running percentage of suspected lies after a 2-star announcement (*i.e.*, the fraction of times a 2-star announcement was followed by the draw of a blank card, indicating a 2/3 probability of being lied to). They include a goodness-of-fit variable, calculated with a Pearson Chi Squared test using, for each individual and each period, the sequence of announcements up to that period and comparing this distribution with the expected distribution given by the objective probability of observing 0, 1, 2 or 3 stars. In the Reputation treatment, the independent variables also include a dummy variable equal to 1 if the project manager is a known liar (he made at least once a 3-star announcement followed by the draw of a blank card up to that period of the game), and an interaction term between this variable and the goodness-of-fit variable.

Both panels confirm that in the Reputation treatment, having been identified as a liar significantly decreases the probability of receiving an investment in the current period: what matters is the undeniable breach of trust that happens when a project manager announces three stars and nature draws a blank card (as opposed to the running number of detected lies). In the No Reputation treatment, few variables are significant after a 2-star announcement. We point out that the running number of detected lies is highly significant and decreases the probability of an investment happening after a 3-star announcement, showing that if they have evidence that the pool of project managers has been dishonest, investors don’t trust a 3-star announcement.

Importantly, the regressions also show that project managers can get away with lies that are harder to detect, as both the coefficients on Goodness of Fit and on the fraction of 2-star announcements with a subsequent loss are largely insignificant. The significant negative time trend observed in both treatments after a 2-star announcement suggests, nevertheless, that investors might become more and more suspicious over time in situations when the announcement may in fact be a deniable lie.

Result 2 summarizes investors’ punishment strategies which provide the reasoning be-

Table 4: Probability of Receiving an Investment After Seeing a 2-star or a 3-star Announcement.

Announcement	Two Stars		Three Stars	
Treatments	No Reputation	Reputation	No Reputation	Reputation
Period	-0.039**	-0.036***	0.044*	0.008
	(0.016)	(0.014)	(0.024)	(0.063)
Goodness of Fit	1.753	1.815	4.571	164.9
	(1.698)	(1.611)	(2.913)	(178.9)
Known Liar	-	-1.337**	-	-2.457 **
		(0.651)		(1.181)
Goodness of Fit x Known Liar	-	4.569	-	-150.8
		(4.139)		(178.4)
Number of Detected Lies	0.007	0.448	-0.395 ***	-0.119
	(0.084)	(0.286)	(0.115)	(0.291)
% 2-star Announcements w/ Loss	-0.661	-0.147	-1.230 *	-2.733
	(0.468)	(0.483)	(0.688)	(2.285)
Constant	1.361***	1.325***	1.794***	3.866**
	(0.375)	(0.405)	(0.528)	(1.901)
Number of observations	527	546	381	178
Number of groups	42	39	42	38

Notes: This Table presents the coefficients from panel Probit regressions of a project manager receiving an investment. Observations are restricted to periods in which the announcement is 2 (first two columns) or 3 (last two columns) stars. "Known Liar" is equal to 1 if the project manager has, up to that point in the game, already been identified as a liar, and 0 otherwise. "Number of Detected Lies" is the running number of times there was a loss after a 3-star announcement. "% 2-star Announcements w/ Loss" is the fraction of times there was a loss after a 2-star announcement up to that point. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

hind the project manager Deniable Lies strategy in the Reputation treatment:

Result 2: Project Managers in fixed partnerships are punished after a detected lie, consistent with our conjecture regarding investor behavior. Investors react to lie detection more than to a bad outcomes per se.

4.3 Monetary and Non-Monetary Motivations for Project Managers' Lies

We now demonstrate that the project managers' strategies are driven by the appetite for personal profit rather than by paternalistic motives or efficiency concerns.

4.3.1 Project Managers' Profits

We start by reporting average earnings by announcement as well as overall for both treatments. This is shown in the top panel of Table 5.

This Table delivers four findings. First, as a result of the investors' risk profile, there is a clear cut-off point in profitability in terms of which announcements lead to higher payoffs. It is only for announcements of at least 2 stars that project managers can attract investments (the mean profit for announcing 1 star is only 44.9 tokens in the No Reputation treatment and 54.4 in the Reputation treatment, while it is 185.2 and 225.9, respectively, after a 2-star announcement). Second, a 3-star announcement pays more on average than a 2-star announcement: 232.7 *vs.* 185.2 in the No Reputation treatment ($p = 0.001$) and 279.2 *vs.* 225.9 in the Reputation treatment ($p < 0.001$). Third, reputation increases profits for 2- and 3-star announcements ($p = 0.005$ when comparing average earnings after a 2-star announcement with and without reputation; $p < 0.001$ when comparing average earnings after a 3-star announcement). Finally, project managers do not benefit overall from having to build a reputation (they earn on average 185.6 in the Reputation treatment compared to 193.1 in the No Reputation treatment; $p = 0.636$) because they infrequently announce 3 stars to remain credible.

Table 5: Average Project Managers' Earnings by Announcement.

Treatments	No Reputation	Reputation
Announced 0 Stars	30	30
Announced 1 Star	44.9	54.4
Announced 2 Stars	185.2	225.9
Announced 3 Stars	232.7	279.2
Average earnings	193.1	185.6
Did Not Use Deniable Lie Strategy	204.5 (91.4%)	172.1 (52.6%)
Used Deniable Lie Strategy	190.0 (8.6%)	202.6 (47.4%)

Notes: Average earnings of the project managers who lied at least once based on announcement, on average, and based on whether the Deniable Lie Strategy was used (in parenthesis we indicate the fraction of subjects who use this strategy).

If on average project managers earn similarly across the two treatments, one strategy that yields higher returns in the Reputation treatment is the Deniable Lie strategy. The bottom panel of Table 5 shows that in the Reputation treatment the projects managers who use this strategy earn on average 202.6 tokens compared to 172.1 for those who did not use it ($p = 0.044$), whereas the respective profits are 190 and 204.5 in the No Reputation treatment ($p = 0.572$). Lying in a way that maintains deniability pays more only when one's reputation is at stake.[31]

[31] We report the average profits by announcement, for each block of nine periods, by treatment, in Appendix Table D4. The findings above are observed in each block of periods and the general trend is that average profits per announcement decrease over time in both treatments, as a result of lie detection. We also report how the Deniable Lie Strategy translates into payoffs for both treatments in each block of 9 periods in

4.3.2 Non Monetary Motivations

We explore whether project managers have good intentions when they choose to lie. Indeed, very few investors invest after a 1-star announcement although it is profitable in expectation. By announcing a higher number of stars, a project manager would increase the investor's expected gain. While we show in the next section that investors are, on average, hurt by the asymmetry in information, we can rule out paternalistic motives even more strongly. We do so by focusing on those individuals who do lie and examine their behavior when faced with 0 stars. In this situation, a project manager knows for a fact that Nature will draw a blank card, resulting in a certain negative outcome for an investor who invests. If project managers who lie are acting out of good will, we should not observe any lies when the true number of stars is 0. Further, there should be no differences between treatments. Yet, 85.3% of project managers in the No Reputation treatment and 63.9% in the Reputation treatment also lie after observing 0 stars. They do so frequently (78.3% of the times they observe 0 stars in the No Reputation treatment, 44.0% in the Reputation treatment), showing that this behavior is both commonplace and persistent. Since over 85% of subjects lie although they know for a fact that this will hurt the investor, we can rule out paternalistic motives for a large majority of liars, as such motives are consistent with the data for no more than 15% of liars.

We also reject that preferences for efficiency are guiding announcements. First, subjects with preferences for efficiency should invest in all periods, a behavior that we do not observe. Given the random assignment of the roles, we assume that project managers share the same social preferences as the investors. Second, in the Allocation Game all but one question presented subjects with efficiency trade-offs (see Appendix Table B1).[32] Two-sided tests of proportion show that subjects who do or do not display preferences for efficiency in the Allocation Game are no different in terms of whether they lie in the Announcement Game, or whether they lie when they face 0 stars.[33] In addition, Probit and OLS regressions show that the number of efficient choices in the Allocation Game has no bearing on the probability of having lied at least once, the number of lies, or having lied at least once when 0 stars were observed (see Appendix Table E1).

Finally, regression analyses that also control for other project managers' characteristics (Machiavellian scores or gender) rule out that innate differences in individual characteristics drive behavior in the Announcement Game. No clear pattern emerges regarding the impact of individual characteristics on a subject's probability of having lied at least once, the

Appendix Table D5. With reputation, project managers who used this strategy earned more in the last two blocks compared with those who did not use it.

[32] A caveat is that choices in the Allocation Game may be influenced by the Announcement Game.

[33] The smallest p-value is $p = 0.195$ when looking at whether answering all Allocation Game questions in a way that is consistent with preferences for efficiency impacts the probability of lying.

number of lies, or having lied at least once when seeing 0 stars, or making a detectable or deniable lie (see Appendix Table E2). Thus, we state our third result as follows:

Result 3: (1) Using the Deniable Lie strategy generates higher returns to the project managers who have to build a reputation. (2) A large fraction (at least 85%) of project managers' decisions on how and when to lie is impacted by strategic considerations in order to increase gains. (3) A small fraction of players (at most 15%) who lie may also/instead be motivated by paternalism, not just personal monetary gains. (4) Machiavellian characteristics or differences in preferences towards efficiency do not explain patterns of lying.

4.4 Investors' Investment Patterns, Errors and Earnings

We next turn to the consequences of the project managers' communication on investors. We investigate investors' decisions and earnings, and compare them with two counterfactual situations: investment under truth-telling, provided by the Truthful Announcement Game, and investment without communication, provided by a new robustness treatment.

4.4.1 Investment Patterns

At the aggregate level, investment behavior is not uniform across all periods: investors invest only 60.0% of the time in the No Reputation treatment and 57.6% in the Reputation treatment. Only 2.4% of investors invest in all periods in the baseline (and none in the Reputation treatment), contrary to predictions based on standard preferences.[34]

Table 6 shows how frequently investors invest for a given announcement, and how this relates to behavior under truthful revelation, as shown in the Truthful Announcement Game.[35] When announcements are low (0 or 1 stars) investors generally find them credible and invest at similar rates compared with the Truthful Announcement Game, regardless of the market structure: they almost never invest when the announcement is 0 and rarely invest when it is 1. However, after a 2-star announcement credibility is an issue in both treatments: they invest significantly less than they did in the Truthful Announcement Game, realizing that such announcements can, in fact, be (deniable) lies ($p < 0.001$). After a 3-star announcement, credibility remains an issue in the absence of reputation, while reputation protects

[34]Risk aversion cannot explain this pattern since a subject should behave in only one of two ways: individuals with a "low enough" level of risk aversion should always invest, whereas those with moderate or high risk aversion should never invest. Preferences for efficiency cannot explain these patterns either since investment yields a higher level of efficiency regardless of Nature's draw. Further, Figure C2 in Appendix C shows that investment slightly decreases over time in both treatments.

[35]Note that there are no treatment differences in how subjects behaved in the Truthful Announcement Game ($p > 0.10$ for all pairwise comparisons for each announcement; tests of proportion). This rules out that subjects had different tolerances towards risk across our two treatments.

investment levels compared with the Truthful Game.[36] With reputation, subjects behave no differently than if they knew the 3-star announcement was truthful.

Table 6: Investment in the Announcement and Truthful Announcement Games, by Treatment.

Treatments		No Reputation	Reputation
0 stars announced:	% Announcement Game	0.0% (25)	0.0% (30)
	% Truthful Ann. Game	2.4%	0.0%
1 star announced:	% Announcement Game	11.4% (40)	9.0% (38)
	% Truthful Ann. Game	11.9%	7.7%
2 stars announced:	% Announcement Game	65.1% (42)	72.6% (39)
	% Truthful Ann. Game	92.9%	97.4%
3 stars announced:	% Announcement Game	77.4% (42)	92.3% (39)
	% Truthful Ann. Game	100%	97.4%

Notes: The Table shows the fraction of times investors invest in the Announcement Game on the first line and in the Truthful Announcement Game on the second line (where announcements were truthful by design). The number of subjects who faced a particular announcement is indicated in parentheses.

4.4.2 Investors' Errors and Earnings

There is no difference in investor earnings across treatments, whether in terms of distributions or simple average earnings (the mean is 152.3 and 154.5 tokens in the No Reputation and Reputation treatments, respectively).[37] The earnings patterns do not reflect a simple rule of proportionality of earnings with the investment decisions, as they are affected by investors' mistakes. Type 1 errors describe periods in which investors do not invest, but would if they believed the announcement; Type 2 errors describe periods in which investors invest, but would not have if they knew the truth. To determine whether a choice is a mistake or not, we look at whether the choice in the Announcement Game is consistent with that in the Truthful Announcement Game (that only reflects the investor's preferences).

Table 7 shows that both types of errors are common. In the No Reputation treatment Type 1 and Type 2 errors happen, respectively, in 23.9% and 21.4% of periods. That is, in close to 1 in 4 periods investors miss out on good opportunities (in their eyes), and in more than 1 in 5 periods they are the victims of fraudulent announcements. Conditional fractions are even more staggering: 60.8% of non-investments are missed opportunities and 35.3% of investments are the result of fraudulent announcements in this treatment. Asymmetric information leads to a high fraction of errors across periods, even in the presence

[36] p=0.267 when comparing behavior with reputation with that in the Truthful Announcement Game. Without reputation, investment drops compared to the Truthful Announcement Game ($p < 0.001$).

[37] Appendix Figure C3 displays the cumulative distribution functions of earnings ($p = 0.519$ in a Kolmogorov-Smirnov test; $p = 0.805$ in a rank-sum test.)

Table 7: Type 1 and Type 2 Errors, by Treatment

	Unconditional		Conditional	
Treatments	No Reputation	Reputation	No Reputation	Reputation
Type 1 Errors	23.9%	16.9%	60.8%	39.9%
Type 2 Errors	21.4%	16.5%	35.3%	28.7%
Periods with errors	45.3%	33.4%	-	-

Notes: This Table reports the average fraction of periods an investor faces a given type of error. Type 1 errors occur when investors do not invest while they would if they had believed the announcement; Type 2 errors occur when investors invest while they would not have if they knew the true number of stars. For conditional values, these fractions are based on Type 1 errors occurring in the sole periods where investment does not happen, and on Type 2 errors in the sole periods where investment happens.

of reputation: this fraction exceeds a third in both treatments. As shown in Appendix Table D6, the frequency of Type 1 errors tends to increase over time in both treatments, consistent with a raising distrust by the investors and their permanent exit of the market.[38] The frequency of Type 2 errors is relatively stable over time: if reputation teaches investors to be more suspicious (because they can appreciate the empirical distribution of announcements by the same project manager), the deniability of lies makes it almost impossible for investors to protect themselves against Type 2 errors. Overall, reputation does not help investors reduce errors over time.

Finally, while in principle lying could lead to better outcomes than truth-telling, by motivating risk averse investors to invest when one star is announced, we show in Appendix F that relative to a situation with market truthfulness, average earnings are significantly lower than what they would be if investors knew the truth (166 tokens *vs.* 152.3 in the No Reputation treatment, and 164.5 *vs.* 154.5 in the Reputation treatment; $p = 0.025$ at most). In that same appendix we report the result of an additional treatment, the No Communication treatment, tested on 72 new subjects with fixed matching, in which project managers cannot send messages to the investors. This treatment reveals that earnings are lower when communication is not allowed at all (investors' earnings are equal to 137.4 tokens *vs.* 154.5 in the Reputation treatment; $p = 0.016$): when project managers can communicate, at least "low" announcements are credible and investors are able to avoid some of the bad projects. Result 4 summarizes the analysis of investors' errors and earnings:

Result 4: Investor earnings are no different between the Reputation and No Reputation treatments. In both treatments losses arise from the lack of transparency, and investors do far worse than under truth-telling, but better than when communication is not possible.

[38]Appendix Table D6 displays the average fraction of periods investors make each type of errors, by treatment, for each block of nine periods and for all periods. Errors do not decrease over time.

5 Discussion and Conclusion

In an economic environment where interest rates are very low, savings accounts pay mediocre returns, which motivates more people to invest in risky assets. Given that a large fraction of individuals have low financial literacy, it is all the more important that financial experts provide reliable advice.[39] However, experts often receive sizeable bonuses that depend on their ability to attract investors and not necessarily on the success of their investments,[40] and so may be tempted to misrepresent their portfolio's expected return to potential customers. This type of misconduct is extremely difficult to measure with field data that can only document detected fraud (Egan et al. (2019)) but not the many lies that are deniable.

Our new Announcement Game shows that in the absence of reputation (*i.e.,* when investors cannot "track" advisors), project managers frequently make Extreme lies that will be detected with certainty *and*, assuming the investor invests, harm the investor. Project managers also frequently make lies that might not be detected (High Risk and Low(er) Risk lies), as well as Deniable lies that cannot be detected in any given period. This delivers two important findings: first, it is crucial to consider not only the prevalence of lies but also their nature; second, in a financial setting exaggerated announcements may be commonplace.

While reputation protects investors against Extreme and High Risk lies, it protects them less against fraudulent announcements with a lower risk of detection, and it does not protect them at all against Deniable lies. Project managers do not become more honest across the board when they have to build a reputation, they mainly change the nature of their lies in order to not face the *financial costs* of being caught lying.

How do our stylized laboratory market settings inform the world outside the lab? In financial markets the turnover of advisers within and across firms is usually high, and people can easily resurface with a "new" identity. Egan et al. (2019) show that the labor market partially undoes firm-level discipline, which lessens the cost of losing one's reputation. In our experiment, reputation, that is the traceability of advisers, lowers the amount of fraud. This is because investors can punish detected liars by not investing with them in the next period. In the absence of individual reputation mechanisms, an investor cannot directly punish a detected lie, and the high prevalence of lies over time reveals the limited impact of punishment by investors. Thus, investors tend to "punish the profession" (and themselves at the same time). Overall, the high level of lies we observe in our stylized settings may contribute to explain the mechanism behind the persistent low confidence that people have

[39] Lusardi and Mitchell (2014) reveal that a majority of people have trouble understanding financial notions such as compounding of interest rates, inflation and stock mutual funds with risk diversification. This replicates in France where we conducted our experiment (Arrondel et al., 2013).

[40] This is the case, among others, for mutual fund managers who are rewarded for fund flows even if they do not outperform passive benchmarks. See Jensen (1968) and Berk and Green (2004).

in the financial sector.[41] This is not to say that investors would fare better without advisers: our No Communication treatment shows that when there is no communication, investors earn less.

Our results do not contradict the literature on cheating games that insists on the importance of the intrinsic cost of deception in the decisions of better informed individuals. Simply, in our setting, the expected monetary benefits of lies are high and investment is lucrative in expectation for the investors. This tends to compensate for the intrinsic costs of lying. Future work could replicate our Announcement Game with different parameters so that in theory, investors should invest conditional on a higher number of stars. This would potentially increase the need for lies, but the moral cost of lying would also increase. Similarly, by reducing the expected payoff of investing, the damage caused by a lie on investors would increase when the state of nature is not good enough in expectation to invest. It would be interesting to explore how individuals would solve the dilemma when moral costs are higher, and how it would modify the distribution of the type of lies. It could make detectable lies less likely but it could affect deniable lies to a smaller extent.

We remain cautious about the external validity of our findings since our experiment did not involve real actors in financial markets, and lying in the lab is not illegal or in violation of professional codes. However, Egan et al. (2019) showed a high level of detectable fraud despite actors having a fiduciary duty towards investors. Our experiment reveals a high frequency of deniable lies in all settings, even though getting caught lying is not illegal. This is concerning for what it implies for markets outside the lab: if there are so many deniable lies when lying is not illegal and detectable lies have no long-term consequences, an implication is that such lies may be very prevalent outside the lab as well. A major challenge is to discourage deniable lying. Our findings call for various types of interventions. One intervention consists of increasing the moral cost of lying, for example by introducing compulsory bankers' oaths in companies. Another one consists of implementing incentive schemes that increase financial advisers' personal responsibility and accountability.

In any case, more centralized investigations and sanctions in case of detected fraud are probably needed. As claimed in Greenspan (2007), *"An area in which more rather than less government involvement is needed, in my judgment, is the rooting out of fraud. It is the bane of any market system."* (p.375).

[41] See, for example, https://www.aei.org/research-products/report/public-opinion-10-years-after-the-financial-crash/ retrieved on October 20, 2021.

References

Abdellaoui, M., P. Klibanoff, and L. Placido (2015): "Experiments on Compound Risk in Relation to Simple Risk and to Ambiguity," *Management Science*, 6, 1306–1322.

Abeler, J., D. Nosenzo, and C. Raymond (2019): "Preferences for truth-telling," *Econometrica*, 87, 1115–1153.

Ambuehl, S., B. D. Bernheim, and A. Ockenfels (2021): "What Motivates Paternalism? An Experimental Study," *Amerocan Economic Review*, 111, 787–830.

Anagol, S., S. Cole, and S. Sarkar (2017): "Understanding the Incentives of Commissions Motivated Agents: Theory and Evidence from the Indian Life Insurance Market," *Review of Economics and Statistics*, 99, 1–15.

Arrondel, L., M. Debbich, and F. Savignac (2013): "Financial Literacy and Financial Planning in France," *Numeracy*, 6.

Austen-Smith, D. (1993): "Interested experts and policy advice: multiple referrals under open rule," *Games and Economic Behavior*, 5, 1–43.

Balafoutas, L., A. Beck, R. Kerschbamer, and M. Sutter (2013): "What drives taxi drivers? A field experiment on fraud in a market for credence goods," *Review of Economic Studies*, 80, 876–891.

Balbuzanov, I. (2019): "Lies and consequences: The effect of lie detection on communication outcomes," *International Journal of Game Theory*, 48, 1203–1240.

Battaglini, M. (2002): "Multiple referrals and multidimensional cheap talk," *Econometrica*, 70, 1379–1401.

Beck, A., R. Kerschbamer, J. Qiu, and M. Sutter (2014): "Car mechanics in the lab - Investigating the behavior of real experts on experimental markets for credence goods," *Journal of Economic Behavior & Organization*, 108, 166–173.

Benndorf, V., D. Kubler, and H.-T. Norman (2015): ""Privacy Concerns, Voluntary Disclosure of Information, and Unraveling: An Experiment," *European Economic Review*, 75, 43–59.

Berk, J. B. and R. C. Green (2004): "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 112, 1269–1295.

Bernard, M., J. Fanning, and S. Yuksel (2018): "Finding cooperators: Sorting through repeated interaction," *Journal of Economic Behavior & Organization*, 147, 76–94.

Beyer, A., D. A. Cohen, T. Z. Lys, and B. R. Walther (2010): "The financial reporting environment: Review of the recent literature," *Journal of Accounting and Economics*, 50, 296–343.

Blume, A., D. Dejong, Y. Kim, and G. Sprinkle (1998): "Experimental Evidence on the Evolution of Meaning of Messages in Sender-Receiver Games," *American Economic Review*, 88, 1323–1340.

Blume, A., E. Lai, and W. Lim (2020): "Strategic information transmission: a survey of experiments and theoretical foundations," in *Handbook of Experimental Game Theory*, ed. by M. Capra, R. Croson, M. Rigdon, and T. Rosenblat, Cheltenham: Edward Elgar Publishing.

Bock, O., I. Baetge, and A. Nicklisch (2014): "Hroot: Hamburg registration and organization online tool," *European Economic Review*, 71, 117–120.

Bohnet, I., H. Harmgart, S. Huck, and J.-R. Tyran (2005): "Learning trust," *Journal of the European Economic Association*, 3, 322—-329.

Brandts, J. and G. Charness (2003): "Truth or consequence: An experiment," *Management Science*, 49, 116–130.

Brown, J. and D. Minor (2016): "Misconduct in financial services: Differences across organizations." *Harvard Business School Working Paper*, 16-022.

Cai, H. and J. T.-Y. Wang (2006): "Overcommunication in Strategic Information Transmission Games," *Games and Economic Behavior*, 56, 7–36.

Charness, G. and N. Garoupa (2000): "Reputation, Honesty, and Efficiency with Insider Information: An Experiment," *Journal of Economics & Management Strategy*, 9, 425–451.

Christie, R. and F. Geis (1970): *Studies in Machiavellianism*, New York: Academic Press.

Cohn, A., E. Fehr, and M. A. Maréchal (2014): "Business culture and dishonesty in the banking industry," *Nature*, 516, 86–89.

Cooper, R. and G. Frank (2005): "The Highly Troubled Ethical Environment of the Life Insurance Industry: Has It Changed Significantly from the Last Decade and If so, Why?" *Journal of Business Ethics*, 58, 149–157.

Crawford, V. P. and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 50, 1431–1451.

Dickhaut, J., M. Ledyard, A. Mukherji, and H. Sapra (2003): "Information Management and Valuation: An Experimental Investigation," *Games and Economic Behavior*, 44, 26–53.

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

Dulleck, U., R. Kerschbamer, and M. Sutter (2011): "The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition," *American Economic Review*, 101, 530–559.

Dziuda, W. and C. Salas (2018): "Communication with Detectable Deceit," *Available at SSRN: https://ssrn.com/abstract=3234695 or http://dx.doi.org/10.2139/ssrn.3234695*.

Ederer, F. and E. Fehr (2017): "Deception and Incentives: How Dishonesty Undermines Effort Provision," *Working Paper*.

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

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

Emons, W. (1997): "Credence Goods and Fraudulent Experts," *RAND Journal of Economics*, 28, 107–119.

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

Feltovich, N. (2019): "The interaction between competition and unethical behaviour : experimental evidence," *Experimental Economics*, 22, 101–130.

FinancialStabilityBoard (2017): "Reducing misconduct risks in the financial sector. Progress report to G20 Leaders 2017," *Boston Consulting Group*.

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

Fischbacher, U. and F. Föllmi-Heusi (2013): "Lies in Disguise – An Experimental Study on Cheating," *Journal of the European Economic Association*, 11, 525–547.

Forsythe, R., R. Lundholm, and T. Rietz (1999): "Cheap Talk, Fraud, and Adverse Selection in Financial Markets: Some Experimental Evidence," *Review of Financial Studies*, 12, 481–518.

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

Gibson, R., M. Sohn, C. Tanner, and A. F. Wagner (2017): "Investing in Managerial Honesty," *Swiss Finance Institute Research Paper*, 17–03.

Gneezy, U. (2005): "Deception: The role of consequences," *American Economic Review*, 95, 384–394.

Gneezy, U., A. Kajackaite, and J. Sobel (2018): "Lying Aversion and the Size of the Lie," *American Economic Review*, 108, 419–53.

Graham, C., R. E. Litan, and S. Sukhtankar (2002): "Cooking the Books: The Cost to the Economy," *Policy Brief*, 106.

Grasshoff, G., Z. Mogul, T. Pfuhler, N. Gittfred, C. Wiegand, A. Bohn, and V. Vonhoff (2017): "Global Risk 2017. Staying the course in banking," *Boston Consulting Group*.

Greenspan, A. (1967): *The Assault on Integrity in: Capitalism: The Unknown Ideal*, New York: Signet.

——— (2007): *The age of Turbulences: Adventures in a New World*, New York: Signet.

Grossman, S. (1981): "The informational role of warranties and private disclosure about product quality," *Journal of Law and Economics*, 24, 461–483.

Guiso, L., P. Sapienza, and L. Zingales (2008): "Trusting the stock market," *Journal of Finance*, 63, 2557–2600.

Gurun, U. G., N. Stoffman, and S. E. Yonker (2018): "Trust Busting: The Effect of Fraud on Investor Behavior," *Review of Financial Studies*, 31, 1341–1376.

Hagenbach, J. and E. Perez-Richet (2018): "Communication with evidence in the lab," *Games and Economic Behavior*, 112, 139–165.

Huck, S., G. Luenser, and J.-R. Tyran (2012): "Competition fosters trust," *Games and Economic Behavior*, 76, 195—-209.

Hung, A. A., N. Clancy, J. Dominitz, E. Talley, C. Berrebi, and F. Suvankulov (2008): "Investor and Industry Perspectives on Investment Advisers and Broker-Dealers," *Technical Report, RAND Center for Corporate Ethics and Governance*.

Hurkens, S. and N. Kartik (2009): "Would I lie to you? On social preferences and lying aversion," *Experimental Economics*, 12, 180–192.

Jensen, M. C. (1968): "The Performance of Mutual Funds in the Period 1945-1964," *Journal of Finance*, 23, 389–416.

Jin, G. Z., M. Luca, and D. Martin (2021): "Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure," *American Economic Journal: Microeconomics*, 13, 141–173.

Kajackaite, A. and U. Gneezy (2017): "Incentives and cheating," *Games and Economic Behavior*, 102, 433–444.

Khalmetski, K., B. Rockenbach, and P. Werner (2017): "Evasive Lying in Strategic Communication," *Journal of Public Economics*, 156, 59–72.

Khalmetski, K. and D. Sliwka (2019): "Disguising Lies — Image Concerns and Partial Lying in Cheating Games," *American Economic Journal: Micro*, 11, 79–110.

Krishna, V. and J. Morgan (2001): "Asymmetric Information and Legislative Rules: Some Amendments," *American Political Science Review*, 95, 435–452.

Li, Y. X. and B. C. Schipper (2020): "Strategic Reasoning in Persuasion Games: An Experiment," *SSRN*.

Li, Z., H. Rantakari, and H. Yang (2016): "Competitive cheap talk," *Games and Economic Behavior*, 96, 65—-89.

Lusardi, A. and O. S. Mitchell (2014): "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 52, 5–44.

McGee, A. and H. Yang (2013): "Cheap Talk with Two Senders and Complementary Information," *Games and Economic Behavior*, 79, 181–191.

Milgrom, P. (1981): "Good news and bad news: representation theorems and applications," *Bell Journal of Economics*, 12, 380–391.

Mimra, W., A. Rasch, and C. Waibel (2016): "Price competition and reputation in credence goods markets: Experimental evidence," *Games and Economic Behavior*, 100, 337—-352.

Mullainathan, S., M. Noeth, and A. Schoar (2012): "The Market for Financial Advice: An Audit Study," *NBER Working Paper*, 17929.

Piskorski, T., A. Seru, and J. Witkin (2015): "Asset Quality Misrepresentation by Financial Intermediaries: Evidence from the RMBS Market," *Journal of Finance*, 70, 2635–2678.

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, 1779–1812.

Rantakari, H. (2016): "Soliciting Advice: Active versus Passive Principals," *Journal of Law, Economics, and Organization*, 32, 719–761.

Rasch, A. and C. Waibel (2018): "What Drives Fraud in a Credence Goods Market? Evidence from a Field Study," *Oxford Bulletin of Economics and Statistics*, 80, 605–624.

Roychowdhury, S. (2006): "Earnings Management through Real Activities Manipulation," *Journal of Accounting and Economics*, 42, 335–370.

Sanchez-Pages, S. and M. Vorsatz (2009): "Enjoy the silence: an experiment on truthtelling," *Experimental Economics*, 12, 220–241.

Schmidbauer, E. (2017): "Multi-period competitive cheap talk with highly biased experts," *Games and Economic Behavior*, 102, 240—-254.

Serra-Garcia, M., E. van Damme, and J. Potters (2011): "Hiding an inconvenient truth: Lies and vagueness," *Games and Economic Behavior*, 73, 244–261.

Shalvi, S., J. Dana, M. J. Handgraaf, and C. K. de Dreu (2011): "Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior," *Organizational Behavior and Human Decision Processes*, 115, 181–190.

Shirley, S. (2020): "Crime and Punishment: Corporate Criminal Prosecution and Corporate Governance," *Working Paper*.

Sobel, J. (2020): "Lying and Deception in Games," *Journal of Political Economy*, 128, 907–947.

Sutter, M. (2009): "Deception through telling the truth?! Experimental evidence from individual and teams," *The Economic Journal*, 119, 47–60.

Sánchez-Pagés, S. and M. Vorsatz (2007): "An experimental study of truth-telling in a sender–receiver game," *Games and Economic Behavior*, 61, 86–112.

——— (2009): "Enjoy the Silence: An Experiment on Truth-Telling," *Experimental Economics*, 12, 220–241.

Thöni, C. (2017): "Trust and Cooperation: Survey Evidence and Behavioral Experiments," *Social dilemmas: New perspectives on trust, A. M. Van Lange, B. Rockenbach, T. Yamagishi, eds.*

Turmunkh, U., M. J. van den Assem, and D. van Dolder (2019): "Malleable Lies: Communication and Cooperation in a High Stakes TV Game Show," *Management Science*, 65, 4795–4812.

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

Vespa, E. and A. Wilson (2016): "Communication with multiple senders: an experiment," *Quantitative Economics*, 7, 1—–36.

Villeval, M. C. (2021): "Public goods, norms and cooperation," *Handbook of Experimental Game Theory, M. Capra, R.T.A. Croson, M.L. Rigdon, T.S. Rosenblat, eds*, 184–212.

Wilson, A. J. and E. Vespa (2020): "Information transmission under the shadow of the future: An experiment," *American Economic Journal: Microeconomics*, 12, 75–98.

Wilson, A. J. and H. Wu (2017): "At-will relationships: How an option to walk away affects cooperation and efficiency," *Games and Economic Behavior*, 102, 487–507.

Zhe Jin, G., M. Luca, and D. Martin (2018): "Complex Disclosure," *Harvard Business School Working Paper*, 2018-15.

Zingales, L. (2015): "Presidential address: Does finance benefit society?" *Journal of Finance*, 70, 1327–1363.

ONLINE APPENDIX

A Instructions

Below we present the instructions that the subjects received in our experiment (translated from French).

INTRODUCTION (All treatments)

Welcome to this experiment on decision-making. Please switch off your mobile phone and store it. You are not allowed to communicate with the other participants throughout the experiment, or we must exclude you from the experiment and from payments.

This session consists of four successive parts. The first two parts are linked but the other parts are independent. The amount earned at the end of the session is the sum of your earnings in the different parts. During the session we do not speak in Euros but in tokens. The conversion rate of tokens into Euros is:

100 tokens = 2.5 Euros

At the end of the session, your total earnings will be paid to you in cash and confidentially in a separate room.

All the decisions you will make in the session are anonymous: you will never have to enter your name in the computer.

You have received the instructions for the fist part. You will receive the instructions for each part after completing the previous part.

PART 1 (All treatments)

This part consists in 17 periods. The first 16 periods are practice periods and nothing that you will decide in these periods will count for determining your payoff. The only period that counts for your payoff in this part is the 17th (and last) period. We describe below the rules and the task, but for the 16 practice periods the announced payoffs are hypothetical.

(In the No Information treatment, this paragraph was replaced with this one: This part consists in 18 periods. The first 16 periods are practice periods and nothing that you will decide in these periods will count for determining your payoff. The only periods that count for your payoff in this part are the 17th and the 18th (and last) periods. At the end of the session the program will randomly select period 17 or period 18 and your payoff in the selected period will constitute your payoff for this part. Each of these two periods has 50 chances out of 100 to be drawn. We describe below the rules and the task, but for the 16 practice periods the announced payoffs are hypothetical.)

Description of the task

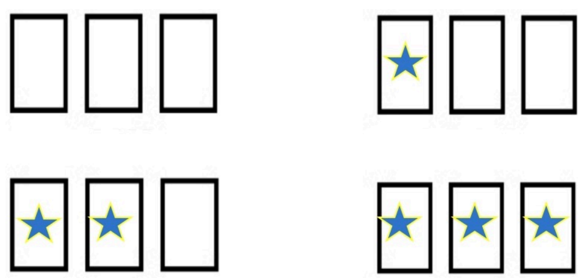
In each period, you receive an initial endowment of 100 tokens and you have to decide between keeping these tokens as the payoff of the period, or to invest them entirely in a project. If this project is a success, you earn three times the number of tokens invested, *i.e.*, 300 tokens. If this project is not a success, you earn 30 tokens.

Description of the investment project

In each period, 3 cards appear on your screen, face down. Each can hide the star symbol or be blank. Each card has 50 chances out of 100 to represent a star and 50 chances out of 100 to be blank. These chances are independent for each card.

Thus, your 3 cards can hide in total no star (which happens with 12.5 chances out of 100), in total 1 star (which happens with 37.5 chances on 100), in total 2 stars (what happens with 37.5 chances on 100), or in total 3 stars (which happens with 12.5 chances out of 100).

You have to press the "Reveal" button to discover your three cards. Depending on the case, the cards will always appear in the following format:



After revealing them, you have to choose whether you invest or not your 100 tokens. After you have made your choice, the program draws one of your three cards, each card having the same chance of being drawn (so, each card has a 1 in 3 chance of being selected).

- If the drawn card represents a star and you have invested, the project is a success and you earn 300 tokens (that is, the endowment of 100 - the investment of 100 + the gain of 300).
- If the drawn card is blank and you have invested, the project is not a success and you earn 30 tokens (that is, the endowment of 100 - the investment of 100 + the winning of 30).

If you have not invested your tokens, you keep your initial endowment of 100 tokens, thus you earn 100 tokens.

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

Remember that in these 16 practice periods, these payoffs are hypothetical and nothing that you decide in these periods counts to determine your actual payoff.

In addition, to ensure that you have met all possible scenarios, we have previously selected cases that correspond to the probabilities announced in the task description above.

Period 17

Period 17 determines your actual payoff for this part. The rules and the task are the same as in the previous 16 periods. The only difference is in the way you have to make your investment choice.

(*In the No Communication treatment, the first sentence was replaced with this one: If it is randomly selected at the end of the session, period 17 determines your actual payoff for this part.)*

Your screen will reveal three cards, face down. Each card has 50 chances out of 100 to hide a star.

In this period, you have to make a decision in each of the following four scenarios. Would you invest or not in the project if the program revealed 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 payoff

Your payoff is determined by the answer to the applicable scenario, that is, the one that corresponds to the total number of stars among your three cards. For example, suppose the three cards hide a total of two stars; in this case, it is your decision in scenario (c) that applies. Another example: suppose the three cards hide a total of three stars; in this case, it is your decision in scenario (d) that applies.

The program then draws one of your three cards.

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

As you can see, the principle is the same as the one that applies in the 16 practice periods, but here you make a decision in every possible scenario.

Since only one of your answers will count to determine your payoff, when you make your decision in each scenario it is in your interest to treat each scenario as if it were the one that actually counts for determining your payoff in this part. You will be informed of the draw and your payoff in this part at the end of the session.

(In the No Communication treatment, the last paragraph was replaced with this one: Since only one of your answers will count to determine your payoff if this period is randomly selected for payment, when you make your decision in each scenario it is in your interest to treat each scenario as if it were the one that actually counts for determining your payoff in this part.)

(In the No Communication treatment only) **Period 18**
If it is randomly selected at the end of the session, period 18 determines your actual payoff for this part. Like in the previous periods your screen will reveal three cards, face down. Each card has 50 chances out of 100 to hide a star. You have to decide again about whether you invest or not in the project. However, in contrast with the previous periods, you have to make a single decision without being informed on the number of cards with a star. It is only at the end of the session that the program will inform you of the total number of stars among your three cards.

Your payoff
If this period 18 is randomly drawn for payment, your payoff is determined as follows:

- *If you have invested in the project and the drawn card is a star, then you earn 300 tokens.*
- *If you have invested in the project and the draw card 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 selected (17 or 18), the card drawn and your payoff in this part at the end of the session.

(All treatments) Please read again these instructions carefully. Whenever you have a question, please raise your hand or press the red button on the side of your desk. We will come immediately to your desk and answer to your question in private.

PART 2 (No Reputation and Reputation treatments)

In this part, each of you will be assigned a role, either A or B. Half of the participants have role A and the other half have role B. Your role remains the same throughout part 2: you will never change role.

Part 2 has a **minimum** of 20 periods and a **maximum** of 40 periods. The exact number of periods was determined before the start of the session.

In each period, each of you is paired so that there is one participant A and one participant B in each pair. You will never know the identity of the people with whom you will be matched. At the beginning of each period, **you are re-paired with a new participant**. Given the number of participants in this session, it is highly unlikely that you will be paired with the same other participant several times in a row.

(*In the Reputation treatment the previous last two sentences were replaced with these ones:* ***Your pair remains the same during all the periods: you interact with the same participant throughout this part.***)

Your task in each period

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

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

Participant B: Participant B receives an initial endowment of 100 tokens and is informed of the announcement of the participant A with whom s/he is matched on his/her number of stars. Participant B has to decide if s/he is willing to keep his/her tokens or invest them in participant A's project.

Determination of payoffs

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

If you are a participant B: Once you have made your investment decision, the program draws one of the three cards of the participant A.

- If you have invested in the project and the drawn card represents a star, the project is a success and you earn 300 tokens (that is, endowment of 100 - investment of 100 + gain of 300).
- If you have invested in the project and the drawn card is blank, the project is not a success and you earn 30 tokens (that is, endowment of 100 - investment of 100 + gain of 30).
- If you did not invest, you earn 100 tokens from your initial endowment.

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

Your screen

On your screen will be a history box where you can see what happened in previous periods. Specifically, you will see four types of information:

(1) your announcements in the previous periods, if you are a participant A, or the announcements of the different participants A with whom you were matched, if you are a participant B;

(2) whether the card drawn in the previous periods represented a star or not, regardless of your role and your decision;

(3) if you have invested in previous periods, if you are a participant B, or if the participants B with whom you were matched have invested or not, if you are a participant A.

(4) your potential payoff in each previous period.

(*In the Reputation treatment the previous sentences were replaced with these ones:*

(1) your announcements in the previous periods, if you are a participant A, or the announcements of the participant A, if you are a participant B;

(2) whether the card drawn in the previous periods represented a star or not, regardless of your role and your decision;

(3) if you have invested in previous periods, if you are a participant B, or if the participant B has invested or not, if you are a participant A.

(4) your potential payoff in each previous period.)

A scroll bar will allow you to scroll through all previous periods.

For each of you, only one period has already been drawn by the program for payment and it is your decision in this period that will help determine your payoff for this part. Thus, it is in your interest to make your decisions in each period as if it was that period that counted to determine your payoff of the part.

Remember that pairs are rematched in each new period.
(*In the Reputation treatment the previous sentence was replaced with this one: Remember that pairs are fixed for the entire part.*)

Please read again these instructions. Whenever you have a question, please raise your hand or press the red button. We will immediately answer to your question in private.

PART 2 (No Communication Reputation treatment)

In this part, each of you will be assigned a role, either A or B. Half of the participants have role A and the other half have role B. Your role remains the same throughout part 2: you will never change role.

Part 2 has a **minimum** of 20 periods and a **maximum** of 40 periods. The exact number of periods was determined before the start of the session.

In each period, each of you is paired so that there is one participant A and one participant B in each pair. You will never know the identity of the people with whom you will be matched. **Your pair remains the same during all the periods: you interact with the same participant throughout this part.**)

Your task in each period

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

Participant A has no task to perform. Participant B cannot see participant A's cards at any time.

Participant B: Participant B receives an initial endowment of 100 tokens and s/he has to decide if s/he is willing to keep his/her tokens or invest them in participant A's project. S/He is not informed of the number of stars on the cards of the participant A with whom s/he is matched.

Determination of payoffs

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

If you are a participant B: Once you have made your investment decision, the program draws one of the three cards of the participant A.

- If you have invested in the project and the drawn card represents a star, the project is a success and you earn 300 tokens (that is, endowment of 100 - investment of 100 + gain of 300).
- If you have invested in the project and the drawn card is blank, the project is not a success and you earn 30 tokens (that is, endowment of 100 - investment of 100 + gain of 30).
- If you did not invest, you earn 100 tokens from your initial endowment.

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

Your screen

On your screen will be a history box where you can see what happened in previous periods. Specifically, you will see four types of information:

(1) whether the card drawn in the previous periods represented a star or not, regardless of your role and your decision;

(2) if you have invested in previous periods, if you are a participant B, or if the participant B has invested or not, if you are a participant A.

(3) your potential payoff in each previous period.
A scroll bar will allow you to scroll through all previous periods.

For each of you, only one period has already been drawn by the program for payment and it is your decision in this period that will help determine your payoff for this part. Thus, it is in your best interest to make your decisions in each period as if it was that period that counted to determine your payoff of the part.

Remember that pairs are fixed for the entire part.)

Please read again these instructions. Whenever you have a question, please raise your hand or press the red button. We will immediately answer to your question in private.

PART 3 (No Reputation, Reputation and No Communication treatments)

This part is independent of the previous parts.

In this part, there are two roles: participant X and participant Y. There is the same number of participants X and participants Y. The part consists of 15 periods.

At the beginning of each period, the program randomly matches each participant X with a new participant Y. Only participant X makes decisions that determine both his/her payoff and that of participant Y.

(*In the Reputation and No Communication treatments the previous paragraph was replaced with this one: At the beginning of the first period and for the whole part, the program randomly matches each participant X with a participant Y. Only participant X makes decisions that determine both his/her payoff and that of participant Y.*)

At the beginning of the part, the program randomly assigns you one of two roles for the entire part, regardless of your role in the previous part. However, you will only be informed of your role at the end of the session. During the part, **you will all make decisions in the role of a participant X**.

If at the end of the session you learn that the program has assigned you the role of a participant Y, none of the decisions you have made will count. Your decisions will count only if the program has assigned you the role of a participant X.

In each period, in the role of participant X you have to choose between two payoff options for you and for participant Y. Here is an example of choices that are presented to you (the actual choices are different from this example):

Option 1: (50, 100) Option 2: (100, 80)

The first number in parentheses is always participant X's payoff and the second number is always participant Y's payoff. In the example above, assuming the period is drawn for payment, option 1 pays you 50 tokens and 100 tokens for participant Y; option 2 pays you 100 tokens and 80 tokens for participant Y.

In each period, you make the same type of decision. Once you have made your decisions in all periods, you will be informed about the period previously drawn for payment by the program. Given your actual role, it is your decision or that of the other participant with whom you are matched in this period that will determine your payoff.

Please read again these instructions. Whenever you have a question, please raise your hand or press the red button. We will immediately answer to your question in private.

PART 4 (No Reputation, Reputation and No Communication treatments)

You will see 20 statements about personality characteristics on your screen. Please indicate for each statement how much you agree or disagree with these statements. Choose the number that corresponds the most to your opinion.

1. Strongly Disagree
2. Disagree
3. Slightly Disagree
4. Neutral
5. Slightly Agree
6. Agree
7. Strongly Agree

- Never tell anyone the real reason you did something, unless it is useful to do so.
- The best way to handle people is to tell them what they want to hear.
- One should take action only when sure it is morally right.
- Most people are basically good and kind.
- It is safest to assume that all people have a vicious streak, and it will come out when they are given a chance.
- Honesty is the best policy in all cases.
- There is no excuse for lying to someone else.
- Generally speaking, people won't work hard unless they are forced to do so.
- All in all, it is better to be humble and honest than to be important and dishonest.

- When you ask someone to do something for you, it is best to give the real reasons for wanting it rather than giving reasons which carry more weight.
- Most people who get ahead in the world lead clean, moral lives.
- Anyone who completely trusts anyone else is asking for trouble.
- The biggest difference between most criminals and other people is that the criminals are stupid enough to get caught.
- Most people are brave.
- It is wise to flatter important people.
- It is possible to be good in all respects.
- P.T. Barnum was wrong when he said that there is a sucker born every minute.
- It is hard to get ahead without cutting corners here and there.
- People suffering from incurable diseases should have the choice of being put painlessly to death.
- Most people forget more easily the death of their parents than the loss of their property.

B The Allocation Game

Table B1: The Allocation Game

Decision	Option 1	Option 2
1	(30,100)	(100,30)
2	(0,0)	(40,30)
3	(30,30)	(60,60)
4	(80,60)	(70,100)
5	(100,30)	(30,230)
6	(70,60)	(90,80)
7	(100,30)	(300,230)
8	(70,60)	(90,50)
9	(30,30)	(30,300)
10	(60,50)	(90,40)
11	(100,30)	(30,230)
12	(230,30)	(100,230)
13	(60,60)	(50,20)
14	(230,230)	(300,230)
15	(60,70)	(50,20)

Note: In parentheses, the first number is the payoff of X and the second number the payoff of Y.

C Appendix Figures

Figure C1 displays the evolution of the relative frequency of lies, by treatment. Figure C2 displays the evolution of the proportion of investors who invest over time, by treatment.

Figure C1: Evolution of the Relative Frequency of Lies, by Treatment

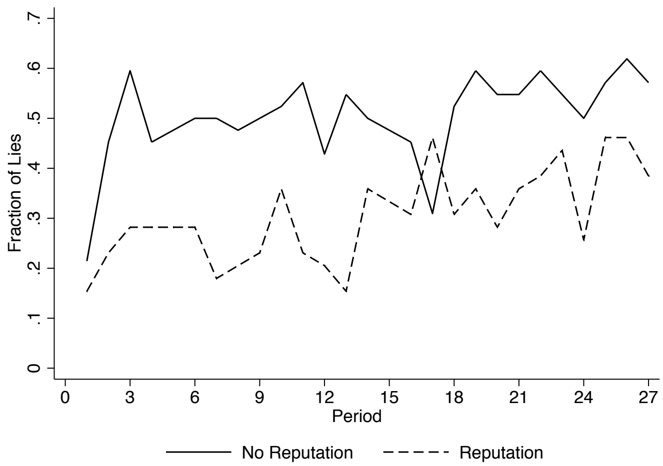


Figure C2: Evolution of the Relative Frequency of Investment, by Treatment.

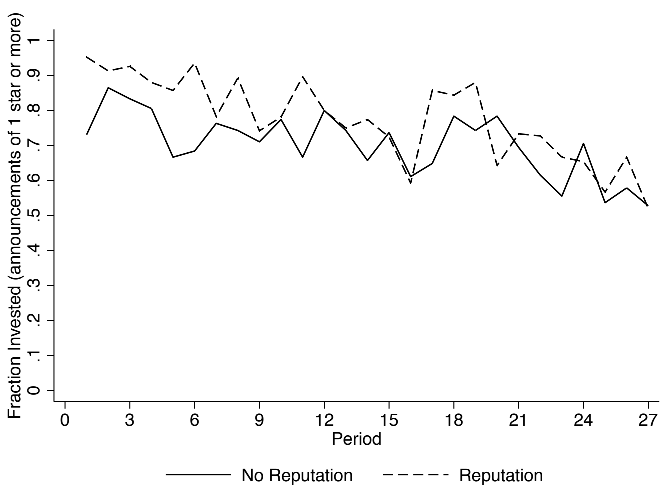


Figure C3 represents the CDF of investors' earnings in the No Reputation, Reputation and No Communication treatments of the Announcement Game, and in the Truthful Announcement Game.

To compute a CDF in the Truthful Announcement Game we pooled the decisions made in the first part of the No Reputation, Reputation and No Communication treatments because this first part was similar across treatments. We then randomly drew 12.5% of observations with 0 or 3 stars each and 37.5% of observations with 1 or 2 stars each. We selected the decisions made in these encounters and then randomly drew one card to determine the simulated payoffs.

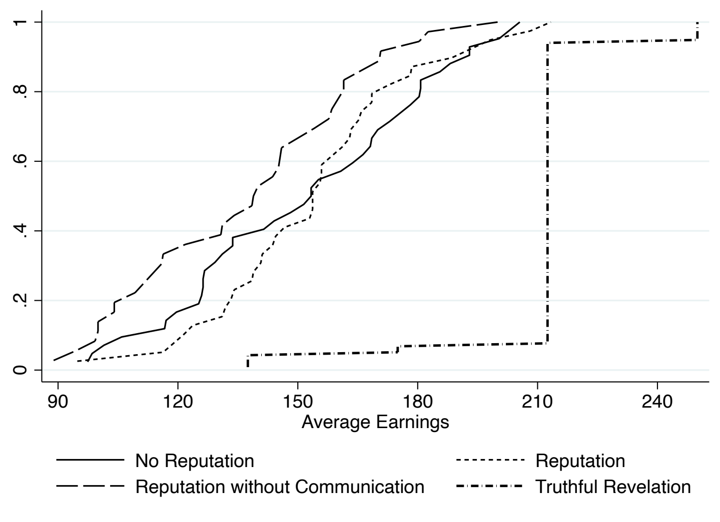


Figure C3: CDF of Investors' Earnings Across Treatments and Under Truthful Revelation.

D Appendix Tables

Table D1 displays the fraction of subjects who engage in each type of lies and the prevalence of lies in the whole sample. It is similar to Tables 2 and 3, except that here we take all the data into account (as opposed to restricting it to project managers who lie at least once). The patterns of data as well as conclusions of the main text apply here as well.

Table D1: A Typology of Lies – Fraction of Subjects who Engage in Each Type of Lies and Prevalence of Lies (Whole Sample)

Treatments	No Reputation	Reputation
Fraction of project managers who engage in each of lie at least once		
Extreme Lies	39.5%	2.7%
High Risk Lies	52.1%	12.8%
Low(er) Risk Lies	64.3%	48.7%
Deniable Lies	81.4%	94.9%
Detectable Lies	69.0%	48.7%
Relative frequency of each type of lie		
Extreme Lies	25.0%	1.4%
High Risk Lies	25.6%	2.5%
Low(er) Risk Lies	34.6%	16.3%
Deniable Lies	43.4%	45.5%

Notes: The top panel displays the fraction of subjects who engage in each type of lie at least once. For example, in the No Reputation treatment 39.5% of the project managers make at least one Extreme lie. The last row indicates the fraction of subjects who make Detectable lies (Extreme, High Risk, and Low(er) Risk lies) at least once. The total of each column is not equal to 100% because each project manager may have used several categories of lies at least once. The bottom panel displays the fraction of times subjects make each type of lies in each treatment, conditioned on the true observed number of stars. For example, in the No Reputation treatment subjects make Extreme lies 25% of the time when they observed 0 stars. The total of each column is not equal to 100% because the prevalence of each category of lie is conditioned on the true observed number of stars. For example, Extreme lies only refer to cases in which project managers have observed 0 stars, while Low(er) Risk lies refer to situations in which they have observed 0, 1 or 2 stars.

Table D2 shows the evolution of Extreme, High Risk, Low(er) Risk and Deniable lies over the first 9, middle 9 and last 9 periods across the two treatments.

Table D2: A Typology of Lies – Fraction of Project Managers who Engage in Each Type of Lies at Least Once

Treatments	No Reputation	Reputation
Extreme Lies		
First 9 Periods	33.3%	0%
Middle 9 Periods	36.8%	0%
Last 9 Periods	33.3%	3.3%
All Periods	*44.1%*	*2.8%*
High Risk Lies		
First 9 Periods	48.6%	5.3%
Middle 9 Periods	42.9%	5.3%
Last 9 Periods	37.1%	10.5%
All Periods	*62.9%*	*13.2%*
Low(er) Risk Lies		
First 9 Periods	60.0%	28.9%
Middle 9 Periods	51.4%	36.8%
Last 9 Periods	62.9%	34.2%
All Periods	*77.1%*	*50%*
All Detectable Lies		
First 9 Periods	71.4%	31.6%
Middle 9 Periods	62.8%	39.5%
Last 9 Periods	62.9%	34.2%
All Periods	*82.9%*	*50%*
Deniable Lies		
First 9 Periods	74.3%	76.3%
Middle 9 Periods	77.1%	89.5%
Last 9 Periods	65.7%	89.5%
All Periods	*85.7%*	*97.4%*

Notes: These data focus on subjects who lie at least once over the course of the Announcement Game. The Table displays the fraction of subjects who engage in each type of lies at least once, broken down by blocks of 9 periods. For example, in the No Reputation treatment 33.3% of the project managers make at least one Extreme lie (observing 0 stars but reporting 3) in the first 9 periods of the game. The total of each column within a block is not equal to 100% because each project manager may use several categories of lies at least once.

Table D3 shows the evolution of the frequency of Extreme, High Risk, Low(er) Risk and Deniable lies over the first 9, middle 9 and last 9 periods across the two treatments.

Table D3: A Typology of Lies – Relative Frequency of Each Type of Lie

Treatments	No Reputation	Reputation
Extreme Lies		
First 9 Periods	25.9%	0%
Middle 9 Periods	29.0%	0%
Last 9 Periods	32.3%	2.5%
All Periods	*27.9%*	*1.4%*
High Risk Lies		
First 9 Periods	29.1%	2.2%
Middle 9 Periods	26.6%	2.6%
Last 9 Periods	32.1%	2.9%
All Periods	*30.7%*	*2.5%*
Low(er) Risk Lies		
First 9 Periods	42.8%	11.3%
Middle 9 Periods	39.2%	17.1%
Last 9 Periods	50.4%	20.4%
All Periods	*41.5%*	*16.7%*
All Detectable Lies		
First 9 Periods	35.1%	6.4%
Middle 9 Periods	32.5%	9.7%
Last 9 Periods	39.3%	10.0%
All Periods	*35.7%*	*8.8%*
Deniable Lies		
First 9 Periods	47.2%	35.9%
Middle 9 Periods	57.2%	48.0%
Last 9 Periods	58.0%	55.2%
All Periods	*52.1%*	*46.7%*

Notes: These data focus on subjects who lie at least once. The Table displays the frequency by which subjects engage in each type of lie. For example, in the No Reputation treatment, in the first block of 9 periods, on average, subjects engage in deniable lies 47.2% of the time. The total of each column is not equal to 100% because the prevalence of each category of lie is conditioned on the true observed number of stars. For example, Extreme lies only refer to cases in which project managers have observed 0 stars, while Low(er) Risk lies refer to situations in which they have observed 0, 1 or 2 stars.

Table D4 shows the evolution of the project managers' average earnings over the first 9, middle 9 and last 9 periods across the two treatments and for each announcement.

Table D4: Average Project Managers' Earnings by Announcement.

Treatments	No Reputation	Reputation
Announced 0 Stars		
First 9 Periods	30	30
Middle 9 Periods	30	30
Last 9 periods	30	30
All Periods	*30*	*30*
Announced 1 Star		
First 9 Periods	56.6	44.0
Middle 9 Periods	54.8	51.8
Last 9 periods	50.8	49.1
All Periods	*44.9*	*54.4*
Announced 2 Stars		
First 9 Periods	201.9	255.8
Middle 9 Periods	197.5	224.2
Last 9 periods	177.5	202.4
All Periods	*185.2*	*225.9*
Announced 3 Stars		
First 9 Periods	252.2	290.4
Middle 9 Periods	240.1	275.3
Last 9 periods	225.4	259.0
All Periods	*232.7*	*279.2*

Notes: Average earnings of the project managers who lied at least once in the first, second or last block of 9 periods, and for the entire game.

Table D5 displays the evolution of the project managers' average earnings over the first 9, middle 9 and last 9 periods, depending on the use of the Deniable Lie strategy.

Table D5: Average project managers' earnings by Deniable Strategy, Treatment and Blocks of Periods

Treatments	No Reputation	Reputation
Did Not Use Deniable Lie Strategy		
First 9 Periods	200.4	197.5
Middle 9 Periods	203.2	164.0
Last 9 periods	196.4	154.6
All Periods	*204.5*	*172.1*
Used Deniable Lie Strategy		
First 9 Periods	234.0	195.0
Middle 9 Periods	207.7	215.2
Last 9 periods	182.3	180.0
All Periods	*190.0*	*202.6*

Notes: Average earnings of project managers, either in the entire game, or by first, second or last block of 9 periods, conditional on the use of the Deniable Lie strategy.

Table D6: Type 1 and Type 2 Errors, by Treatment

	Unconditional		Conditional	
Treatments	No Reputation	Reputation	No Reputation	Reputation
Type 1 Errors				
First 9 Periods	19.1%	9.7%	53.3%	24.6%
Middle 9 Periods	21.7%	16.2%	55.8%	41.3%
Last 9 Periods	31.0%	24.8%	71.3%	51.2%
All Periods	*23.9%*	*16.9%*	*60.8%*	*39.9%*
Type 2 Errors				
First 9 Periods	22.0%	16.2%	34.2%	26.8%
Middle 9 Periods	21.7%	15.4%	35.5%	25.4%
Last 9 Periods	20.6%	18.0%	36.5%	34.8%
All Periods	*21.4%*	*16.5%*	*35.3%*	*28.7%*
Periods with errors				
First 9 Periods	41.0%	25.9%	-	-
Middle 9 Periods	43.4%	31.6%	-	-
Last 9 Periods	51.6%	42.7%	-	-
All Periods	*45.3%*	*33.4%*	-	-

Notes: This Table reports the average fraction of periods an investor faces a given type of error. Type 1 errors occur when investors do not invest while they would if they had believed the announcement; Type 2 errors occur when investors invest while they would not have if they knew the true number of stars. For conditional values, these fractions are based on Type 1 errors occurring in the sole periods where investment does not happen, and on Type 2 errors in the sole periods where investment happens.

E Additional Regression Analyses

The regressions in Table E1 support our investigation of the role of preferences for efficiency, as determined by choices in the Allocation Game, on lying behavior in the Announcement Game. They report the estimates of coefficients from Probit models (models (1) and (3)), and from OLS model (model (2)). In models (1) and (2), standard errors are clustered at the individual level, and in model (3) they are clustered at the session level. The dependent variable is having lied at least once in the Announcement Game in model (1), the number of lies made by the project manager in model (2), and having lied at least once when observing 0 stars in model (3). In the three models, the independent variable is the number of efficient choices made by the project manager in the Allocation Game.

Table E1: Preference for Efficiency in the Allocation Game and Lying in the Announcement Game

	Having Lied At Least Once (1)		Number of Lies (2)		Having Lied At Least Once When Observing 0 Stars (3)	
Treatments	No Rep	Rep	No Rep	Rep	No Rep	Rep
Number of efficient choices	0.126	0.027	0.860	-0.173	0.048	-0.058
	(0.134)	(0.025)	(0.607)	(0.400)	(0.129)	(0.118)
Constant	-0.519	0.654**	3.300	10.254*	0.146	0.986
	(1.585)	(0.322)	(7.485)	(5.107)	(1.556)	(0.402)
Number of observations	42	39	42	39	38	37
R^2	-	-	0.038	0.005	-	-

Notes: This Table presents the coefficients from Probit regressions in models (1) and (3). Standard errors are clustered at the individual level in models (1) and (2) and at the session level in model (3). The dependent variable is having lied at least once in the Announcement Game in model (1), the number of lies made by the project manager in model (2), and having lied at least once when observing 0 stars in model (3). In model (1), the independent variable is omitted for the Reputation treatment because a number of efficient choices not equal to 8 predicts success perfectly - we therefore present the results of the OLS regression instead. $^{**}p < 0.05$, $^{*}p < 0.1$

The regressions in Table E2 confirm the results from the non-parametric tests we report in the main body of the paper about the project managers' behavior. This Table reports the estimates of coefficients from Probit models with standard errors clustered at the individual level. In models (1) and (2) the dependent variable is making a detectable lie (Extreme, High Risk and Low(er) Risk lies). In models (3) and (4) the dependent variable is making a deniable lie. The independent variables include the Reputation treatment, with the No Reputation treatment as the omitted category, as well as a time trend. Models (1) and (3) also control for individual characteristics (number of efficient choices in the Allocation Game, Machiavellian score and gender). These regressions confirm that detectable lies are less frequent in the presence of reputation (regressions 1 and 2), but there are no treatment differences when it comes to Deniable lies (models 3 and 4). This is the case whether or not we use controls. While not reported in detail here, none of the controls are significant.

Table E2: Treatment Differences in the Likelihood of Making Detectable and Deniable Lies.

	Detectable (Model 1)	Detectable (Model 2)	Deniable (Model 3)	Deniable (Model 4)
Reputation	-1.266 (0.276)***	-1.264 (0.273)***	-0.081 (0.137)	-0.052 (0.134)
Period	0.011 (0.008)	0.011 (0.008)	0.017 (0.005)***	0.016 (0.005)***
Mach IV Tactic	-0.007 (0.016)	-	0.003 (0.008)	-
Mach IV Cynism	0.016 (0.016)	-	0.006 (0.008)	-
Mach IV Morality	0.019 (0.079)	-	-0.005 (0.042)	-
Female	-0.180 (0.302)	-	0.102 (0.146)	-
Number Efficient Choices	0.052 (0.073)	-	0.018 (0.033)	-
Constant	-1.393 (0.811)*	-0.871 (0.218)***	-1.380 (0.429)***	-0.982 (0.138)***
Number of observations	1,971	1,971	1,971	1,971
Number of groups	73	73	73	73

Notes: This Table presents the coefficients from panel Probit regressions of making a detectable lie (first two models) or a deniable one (last two models) on treatment dummies. The omitted dummy is the No Reputation treatment. In the first and third models, controls include the number of efficient choices in the Allocation Game, the Machiavellian scores as well as gender. The Machiavellian scores correspond to three dimensions of Machiavellianism: interpersonal tactics (*e.g.*, "It is wise to flatter important people."), cynical view of human nature (*e.g.*, "The biggest difference between most criminals and other people is that criminals are stupid enough to get caught"), and utilitarian morality (*e.g.*, "People suffering from incurable diseases should have the choice of being put painlessly to death"). Standard errors are clustered at the individual level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

F Earnings Counterfactuals

F.1 Investors' Earnings Under Truth-Telling

How do investors' earnings compare to the earnings that would exist under truth-telling? In principle, lying could lead to better outcomes than truth-telling. Indeed, in the Truthful Announcement Game few investors invest when 1 star was announced, showing risk aversion. In expectation, however, they would be better off financially if they did invest at 1 star, which they might have done if project managers persuaded them that they had two stars. But in fact, relative to a situation with market truthfulness, average earnings are significantly lower than what they would be if investors knew the truth (166 *vs.* 152.3 in the No Reputation treatment, and 164.5 *vs.* 154.5 in the Reputation treatment).[42,43] Overall, while investors may be pushed to invest more at 1 star than they otherwise would, the gains from these investments are far outweighed by the losses that come from investment errors that the asymmetric information generates.

F.2 Investors' Earnings in the Absence of Communication

Another interesting counterfactual is the investment behavior and earnings that would exist in the absence of communication. In a new treatment we removed communication in the Reputation treatment and measured the extent to which communication improves investors' earnings despite project manager biases. We recruited an additional 72 subjects from the same subject pool as in the main treatments for four sessions of the **No Communication treatment**. This treatment introduces two changes relative to the Reputation treatment. First, in each period of the Announcement Game, project managers still observe the three cards they are dealt but they take no action, and the investor has to make a decision as to whether invest or not without facing any announcements. Second, in part 1 subjects play an 18th period where they have to decide whether to invest or not without knowing the content of the three cards.

The results show that the primary driver of investment absent communication is risk preferences. Despite similar investment frequencies (57.6% and 56.9% in the treatments with and without communication, respectively, $p = 0.675$), earnings are substantially lower in the No Communication treatment: 137.4 tokens *vs.* 154.5 in the Reputation treatment ($p = 0.016$). Indeed, absent communication, investors often invest when it is not profitable, whereas when project managers can communicate, "low" announcements are credible and investors are able to avoid some of the bad projects.

[42]To compute these hypothetical earnings we considered what the subject would have chosen had he or she known the truth (based on the preferences revealed in part 1); based on this hypothetical choice and the probability of a star being drawn given the true number of stars, we determined the payoff the subject would have had in expectation if they knew the truth. Signed-rank tests comparing average earnings to the earnings if investors knew the truth are at most $p = 0.025$.

[43]The difference is very obvious at the distribution level as well as can be seen when we compare the CDF of payoffs in each treatment of the Announcement Game with the CDF of earnings in the Truthful Announcement Game (see Appendix Figure C3).