



Título artículo / Título **The role of ex post transparency in information transmission -An experiment**
article:

Autores / Autors

Sascha Behnk, Iván José Barreda Tarazona, Aurora García Gallego

Revista:

Journal of Economic Behavior & Organization, 2014, 101: 45-64

Versión / Versió:

PDF Preprint Autors

Cita bibliográfica / Cita bibliogràfica (ISO 690):

BEHNK, Sascha; BARREDA-TARRAZONA, Iván; GARCÍA-GALLEGO, Aurora. The role of ex post transparency in information transmission—An experiment. *Journal of Economic Behavior & Organization*, 2014, 101: 45-64.

url Repositori UJI:

<http://doi:0167-2681 10.1016/j.jebo.2014.02.006>

The role of *ex post* transparency in information transmission - An experiment

Abstract: Asymmetric information in economic relationships often provides incentives to deceive. Previous findings show that *ex ante* disclosure of conflicts of interest not only fails to improve these relationships but leads to even more deception. This study proposes that providing *ex post* transparency could play an important role in reducing deception. Different scenarios of a sender-receiver game allow us to show that disclosing conflicts of interest *ex-post* not only does not induce more deception but also that a 50% chance of subsequent disclosure significantly reduces deception in the case in which the sender obtains a small gain at the expense of a comparatively big loss for her counterpart. We find no significant reductions under other circumstances. Regarding the receivers, we find that the increase in honest messaging, when it takes place, is not anticipated.

Key words: information transmission; deception; transparency; social image; experiment

JEL Classification: C91; D03; D63; D82

1. INTRODUCTION

Asymmetric information is a widespread characteristic in economic relationships and often provides incentives to deceive the less informed party. Private investors, for instance, face such a critical situation. On the one hand, they often rely on financial advisors due to the lack of financial literacy and high transaction costs. On the other hand, commissions for selling specific financial products can incentivize advisors to recommend financial products that are not necessarily the best option for their clients and hence turn a relationship that is designed to be beneficial for investors into the opposite.¹ The subprime mortgage crisis has shown that even institutional investors can be involved in barely controllable relationships with agents evaluating investment opportunities for their clients, such as investment banks or rating agencies, while they have misaligned incentives.

Similar relationships between better-informed agents and their principals can be found in nearly all areas of economic activity, e.g., between salesmen and consumers, lawyers and their clients or workers and company owners. This is an important issue not only for moral concerns but also because substantial economic consequences can arise from deceptive behavior. That these consequences can go far beyond the loss of the deceived ones has been shown by Akerlof's (1970) well-known market for lemons.

A question that has raised much interest in the literature is what factors regarding the information transmission process in principal-agent relationships affect the better informed agents' behavior in situations in which monetary incentives to deceive exist. In this study, we investigate a *prima facie* factor in this context: transparency. The aim of transparency would be to reduce deception by providing principals with cost-free information about their agents' incentives to deceive. However, *ex ante* transparency about conflicts of interest does not seem to keep people from deceiving - even the opposite

¹ For a detailed overview of the role of advice and corresponding incentives on the market for retail financial services, see Popova (2010).

has been observed. When conflicts of interest are disclosed from the beginning, people are apparently provided with a “license to lie” (Cain et al., 2005). Based on similar findings, Inderst et al. (2010) conclude that “increasing transparency is not necessarily a good policy advice”.

Our study addresses this conclusion and contributes to the discussion about the role of transparency in the information transmission process by introducing *ex post* disclosure.² To the best of our knowledge, we are the first to experimentally investigate how costless subsequent information affects decision making in different scenarios of a sender-receiver game with anonymous players. We test whether fewer deceptive messages are sent when the sender’s incentive structure will be disclosed with some probability after the game has been played. To gain deeper insight into the effect that transparency has on the information transmission process in our setting, we develop a belief-based theoretical framework and later examine the moderating role of players’ beliefs in an econometric model.

Three treatments are implemented in which subsequent information is the only factor that changes across treatments. Our baseline is a classic sender-receiver game without disclosure in which the sender faces monetary incentives to deceive. In a second treatment, we introduce assured subsequent disclosure; that is, after the decisions have been made, receivers always find out about the payoff structure of the game played. In a third treatment, subsequent transparency is implemented with only 50% probability.

Our results show that the provision of *ex post* transparency never leads to a higher rate of deception compared to the baseline. The so-called “license to lie” effect, which has been observed in studies concerning *ex ante* transparency, does not appear with subsequent disclosure. Compared to the case without disclosure, we find that deceptive messages are significantly lower only when subsequent disclosure is implemented with 50% probability, and a small profit can be gained by the sender at the expense of a relatively large financial harm to her counterpart.

We propose that the effect of *ex post* disclosure is related to people not only caring about their own behavior and its consequences *per se* but also seeking to protect their social image in the sense of a preference for being perceived by others as a person with moral integrity (e.g., Andreoni and Bernheim, 2009). Furthermore, our within-subject investigation reveals different effects of *ex post* disclosure among our payoff scenarios and hence implies that people care about both the mutual monetary consequences of their behavior *and* how this behavior is perceived by others - similar to the social image model developed by Bénabou and Tirole (2006).

An example that illustrates the interaction between selfish motives and social image through *ex post* transparency is to imagine a person who knows the way to the best restaurant in town, but her family owns the second best place to eat. Why would she tell the truth to a foreigner? And why would she be even more likely to do so if she knew that the foreigner will discover the truth *ex post*? The first inclination would certainly correspond to the attraction toward telling the truth beyond selfish interest. This inclination favoring truthful information transmission would be driven by lying and guilt aversion. However, the second force may be the result of her concern for the other’s impression of her and his possible feelings of regret or deception or even his image of the people of her town. In fact, given that our experimental protocol

² A practical application of *ex post* transparency is to use it as a basis for penalties and deterrence. A recent example is the practice of the German Federal Financial Supervisory Authority: The authority established an internal register (“Beraterregister”) of financial advisors in which information about miscounseling and selling sub-optimal products is collected. This *ex post* disclosure of fraud-near behavior to the authority is supposed to prevent deception by punishment and a threat to the social image and reputation of the treacherous advisors, which are important business requirements in the financial sector.

guarantees anonymity, the phenomenon studied here should be the result of the two last motivations: caring for the other's feelings and the beliefs he could form concerning the people surrounding him.

Receivers, on their part, do not significantly anticipate the decrease in deception that a 50% probability of subsequent disclosure can provide. The message acceptance rates do not increase significantly in the treatments with *ex post* transparency compared to the baseline. We find that first-order beliefs about the sender's decision significantly affect the behavior of the receiver in the expected way: the higher the probability assigned by receivers to senders having sent an honest message, the more likely it is that they accept the message.

The paper is structured as follows: Section 2 offers a review of the related literature and the conceptual framework of the study. Section 3 describes the experimental design, treatments and procedures. Section 4 presents the theoretical predictions and hypotheses. The results are presented and discussed in section 5. Section 6 concludes. Experimental instructions and a screenshot of the program used in the experiment can be found in the appendix.

2. LITERATURE REVIEW AND FRAMEWORK

Since the seminal paper by Crawford and Sobel (1982), many recent contributions have addressed the information transmission process in sender-receiver settings. The main results of Crawford and Sobel's (1982) theoretical model support the assumption that the more the players' preferences diverge, the noisier the equilibrium information is, and if "interests diverge by a given 'finite' amount, only no communication is consistent with rational behavior". By modifying this model, Kartik et al. (2007) show that, when communication is costly, transmitted information is inflated and will lead to deception if receivers are credulous.

Dickhaut et al. (1995) experimentally test Crawford and Sobel's model and basically confirm their theoretical predictions by varying the difference of the players' interests in a sender-receiver game (for a different setting, see Blume et al., 2001). However, some studies, such as those by Cai and Wang (2006) and by Ottaviani and Squintani (2006), find an "over-communication effect", i.e., senders gradually transmit more information to the receiver than predicted in the Crawford and Sobel theoretical equilibrium.

The over-communication effect has raised great interest in examining why people do not always act like the selfish *homo economicus* in these settings but also seem to be affected by specific motivations that can keep them from exploiting an asymmetric information distribution. For instance, concerning the rates of individuals who lied in their experiment to gain a higher profit, Lundquist et al. (2009) state that "these figures are clearly below 100%, indicating that at least some subjects find it psychologically costly to lie".³ That is, some people would experience a disutility from lying that exceeds the corresponding expected monetary gains. In Kartik's (2009) theoretical model, the degree of language inflation and how much information is revealed depends on the intensity of a lying cost.

Important insight concerning this topic has been gained from a wide range of studies concerning the punishment of unfair actions. Brandts and Charness (2003) find that people are much more likely to

³ Non-monetary obstacles to lie are also referred to as "costs of lying" (Gneezy, 2005) or "moral costs" (Inderst et al., 2010). Here, we use the term "costs of lying" to name an intrinsic motivation that keeps people from telling a lie in order to gain higher monetary payoffs.

punish an opportunistic action in a sender-receiver game when it is preceded by a deceptive message that is unfavorable to the receiver. Their findings indicate that people's decisions are not only influenced by their expected payoffs but also by the way they can achieve them. Sánchez-Pagés and Vorsatz (2007) investigate the effect of costly punishment in different sender-receiver games and show that players can be partitioned into two groups, one group of players with preferences for truth-telling and another group taking into account only economic incentives. Sánchez-Pagés and Vorsatz (2009) show a significant effect of punishment by using a game setting with several rounds in which players change their roles. A new feature in their sender-receiver game is that senders have the costly possibility to remain silent instead of sending messages. The authors find that remaining silent is positively correlated with truthful behavior. Subjects who punish deception consistently behave more honestly and remain silent more often.

The existing literature also reflects a special interest in knowing under what circumstances people lie to maximize their own profits at the expense of others and when they are reluctant to deceive their counterparts. Gneezy (2005) experimentally examines the role of consequences in a sender-receiver relationship, in which, unlike the receiver, the sender obtains full information about the game. Gneezy finds that changes in their own wealth through lying have an important effect on subjects' behavior, but people are also sensitive to the harm that deception would cause to their counterparts. Hurkens and Kartik (2009) show in a modified setting that Gneezy's results are consistent with the idea that either a person will never lie or a person will lie when she prefers the outcome provided by lying over the payoffs provided by telling the truth. Recent results obtained by Gibson et al. (2013) point to the fact that people's preferences for truthfulness do not seem to be consistent with a type-based model with simply "economic types" or "ethical types" but rather with a continuum across the two extremes. The authors find heterogeneous preferences for lying and truth-telling within individuals.

Two intrinsic motivations that affect decision making in this sense are lying and guilt aversion. According to López-Pérez and Spiegelman (2013), aversion to lying *per se* is independent of the consequences of a lie for other parties, whereas guilt aversion results from negative consequences to other people caused by the lie and is related to the beliefs that the potential liar has about her counterpart's expectations.

Recently, there has been a growing debate in the literature about which of these intrinsic motivations can actually keep people from deception. Battigalli et al. (in press) mathematically derive how a theory of guilt aversion can accommodate the outcome of Gneezy's (2005) design. Some studies find evidence for guilt aversion (e.g., Charness and Dufwenberg, 2006 & 2010, Reuben et al., 2009, Peeters et al., 2012), whereas other findings do not support this type of intrinsic motivation (e.g., Ellingsen et al., 2010). Another group of studies provide evidence consistent with aversion to lying *per se* (Ellingsen and Johannesson, 2004, Vanberg, 2008, Lundquist et al., 2009, Gneezy and Erat, 2012). López-Pérez and Spiegelman (2013) investigate a specific variant of lying aversion that is based on second-order choice expectations but do not find evidence for this kind of belief-dependent lying aversion. Our aim is not to intervene in this discussion; nevertheless, we consider in our theoretical framework that a combination of both lying and guilt aversion might have an effect on decision making, like several authors propose (e.g. Charness and Dufwenberg, 2010, Gneezy and Erat, 2012), and implicitly that preferences for truthfulness are heterogeneous (Gibson et al., 2013).

An important question that arises from the aforementioned discussion is how people can be kept from deceptive behavior. A particular solution was examined by Popova (2010), who investigates the impact of

voluntary and obligatory payments on financial advisory. She uses a similar sender-receiver game to that reported in Gneezy (2005) and adds a third, a Pareto-dominated option, like in Rode (2010), to avoid a strategic behavior known as “deception through telling the truth” (Sutter, 2009). Popova finds that receivers are able to reduce deception by paying an amount of money to the senders as an incentive for honesty. However, this strategy is only successful when payments are made voluntarily.

Our study focuses on a different factor that could possibly affect people’s reluctance to deceive: transparency. A well-examined type of transparency is the *ex ante* disclosure of conflicts of interest between senders and receivers. However, *ex ante* transparency does not necessarily lead to the optimal outcome for all involved parties (Li and Madarász, 2008). In their experiment, Cain et al. (2005) find that *ex ante* transparency even increases deceptive behavior rather than reducing it (see also Koch and Schmidt, 2010). Disclosing conflicts of interest from the beginning seems to provide agents with a “license to lie”. In a follow-up article, Cain et al. (2011) provide deeper insight on how disclosure can backfire and harm those who should benefit from transparency. They show that both strategic behavior for gaining more profit and moral considerations play a role in biasing advice with *ex ante* disclosure. Interestingly, deceptive advice is rated by their subjects as being, on average, “somewhat unethical” when no disclosure is provided and “somewhat ethical” when conflicts of interest have previously been disclosed. Rode (2010) enhances Gneezy’s design with additional payoff options and provides receivers with different degrees of uncertainty. He also finds that *ex ante* disclosure of conflicts of interest “makes advisors less morally bound to send accurate information”. A competitive framing through pre-game tasks does not change the sender’s behavior but leads to less trust in the sent messages. Inderst et al. (2010) also obtain higher rates of deception when testing different degrees of disclosing conflicts of interest. They conclude that “increasing transparency is not necessarily a good policy advice”.

In this study, we test how *ex post* disclosure affects deceptive behavior. We find a significant reduction in deception when the conflict of interest and the sender’s behavior is revealed to the receiver under specific conditions, i.e., behavior is exposed with 50% probability and the sender can obtain only a small gain at the expense of a comparatively great financial harm to the receiver. We propose that the effect of *ex post* disclosure in our experiment is directly related to people’s tendency not only to care about their own behavior per se and its consequences but also about how this behavior is perceived by others, forming their social image.

A large body of literature finds evidence for social image effects in different settings. Concerns for social image and official status are not only addressed through wealth signals such as luxury goods consumption but also through the creation of a generous public impression (e.g., Lacetera and Macis, 2008). By analyzing theoretically the effect of control in principal-agent relationships, Ellingsen and Johannesson (2008) show that altruistic behavior can be driven by social esteem depending on the audience. In particular, fairness motives seem to be an important factor for establishing and maintaining a desirable social image. A variety of experimental studies using dictator games confirm that subjects act less fairly when actions are not observable (e.g., Dana et al, 2007) or social distance increases (e.g., Hoffman et al, 1996). Andreoni and Bernheim (2009) conclude that instead of altruism or fairness motives, a preference for being publicly perceived as fair can explain these social image effects through an audience.

Furthermore, our between-subject results among different payoff-scenarios, i.e., the payoff-dependent effect of *ex post* transparency, are related to existing models about the interference of monetary incentives

and social image effects. Bénabou and Tirole (2006) develop a model in which they account for the idea that "people's actions indeed reflect a variable mix of altruistic motivation, material self-interest and social or self-image concerns". Their model is mainly based on the interference of monetary incentives and charitable giving. Experimental evidence can be found, e.g., in Ariely et al. (2009).

Taking these findings into account, we assume that social image effects play a key role in driving the reduction in deception through *ex post* disclosure in our experiment and particularly that this type of motivation reinforces or complements lying and guilt aversion in the treatments in which a revelation of immoral actions is possible.

3. EXPERIMENTAL DESIGN

3.1 General game structure

We used a sender-receiver game, based on the game structure reported in Gneezy (2005), to experimentally test for the effect of subsequent disclosure on deceptive behavior. The main task of the sender was to send a message to her counterpart, the receiver, who accepted or rejected that message and, by doing so, determined the payoffs for both players. The game was played as a one-shot interaction. Participants did not know the identity of their counterparts at any time during or after the experiment and were randomly matched in pairs by the computer.⁴ After the subjects of each pair were randomly assigned by the computer the neutrally framed roles of "player 1" (the sender) and "player 2" (the receiver), three different scenarios were shown exclusively to the sender. Each of the scenarios contained three different options: A, B and C. Every option consisted of a payoff for the sender and a payoff for the receiver.⁵ The general structure of the options for each of the scenarios, presented to both player types in the experimental instructions (see Appendix B), was as follows:

Option A: Player 1 receives ... euros and player 2 receives ... euros.

Option B: Player 1 receives ... euros and player 2 receives ... euros.

Option C: Player 1 receives ... euros and player 2 receives ... euros.

Both players' payoffs in each option were shown only to the sender for each of the three scenarios. The task of the sender in each scenario was to choose one of the following three messages to send it to the receiver:

Message 1: Option A will earn you more money than the other two options.

Message 2: Option B will earn you more money than the other two options.

Message 3: Option C will earn you more money than the other two options.

⁴ The experiment was programmed in z-Tree (Fischbacher, 2007).

⁵ We presented the scenarios on different screens and randomized the order of their appearance. The appearance of the three options in each of the scenarios was also randomized. Therefore, the three payoff pairs were randomly assigned the labels "option A", "option B" and "option C" so that senders were not able to predict the payoffs of a specific option label.

After the sender chose a message for each of the three scenarios, the computer randomly selected one of the scenarios as the basis for the game. The specific message that the sender had chosen for the selected scenario was presented to the receiver.

Note that the payoffs in each option were private information to the sender. The receiver only knew that there were three options available in each scenario and that her counterpart would send her a message stating which option was the best one for her, but she did not know anything about the payoffs associated with each option.

The receiver then decided whether to “accept” or “reject” the message. In case she accepted it, the option in the message determined the final payoffs for both players. In case of rejection, one of the other two remaining options in the selected scenario was randomly chosen by the computer to determine the players’ final payoffs, as in Popova (2010). The information that was presented to the players on their final screen depended on the treatments, which are described in section 3.4. After the final screen was shown, the game ended.

3.2 Payoff distribution

In Table 1, we present the payoff distribution for options A, B and C in each of the three scenarios. The interests of both players are misaligned between options A and B, i.e., the best option for the sender is not the best one for the receiver.⁶ The payoff distribution for these two options is mainly based on the three scenarios used by Gneezy (2005).⁷

Scenario	Option	Payoff sender	Payoff receiver
1 (low+;low-)	A	5	6
	B	6	5
	C	3	3
2 (low+;high-)	A	5	15
	B	6	5
	C	3	3
3 (high+;high-)	A	5	15
	B	15	5
	C	3	3

Table 1
Payoff distribution for both players by scenario and option (in euros).

Similarly to Rode (2010) and Popova (2010), we added a Pareto-dominated third option to each of the scenarios, providing the same payoffs for both players for two reasons: First, we wanted to control for the strategic “deception by telling the truth” that was examined by Sutter (2009), who observed that some players act contrary to their actual intentions to reach their goals. By adding the third option to each scenario, we avoided the situation in which the sender would tell the truth to actually deceive the receiver, anticipating that she would reject the message. This is so because the computer randomly implemented one of the two remaining options in case the receiver rejected the message. Second, Sánchez-Pagés and

⁶ Appendix A shows a sample screenshot of one of the scenarios that was presented to senders.

⁷ Gneezy (2005) shows that different payoff distributions dramatically change deception rates across individuals. In our case, this is also true for a within-subject design.

Vorsatz (2009) have shown that some subjects prefer to remain silent, even if this is costly, rather than to make a decision between deception and honesty. In this sense, the Pareto-dominated option serves as an additional alternative to truth-telling because choosing this option avoids a choice between monetarily self-harming honesty and other-harming deception at the cost of a minor but equalitarian payoff for both parties. A behavior that is closely related to this intention is inequity aversion, i.e., a preference for an equal payoff distribution even at a cost for all involved parties (see, e.g., Fehr et al., 1999).

Although option C always provides three euros for each player, the intensity of the payoff misalignment between option A and option B changes appreciably among the three scenarios. In scenario 1, the difference between the payoffs that players receive from options A and B is relatively small. Hence, a successful deception in this scenario leads to a comparatively low gain for the sender compared to an equally low loss for the receiver. We label this scenario (low+,low-).⁸ In scenario 2 (low+,high-), the sender is again able to achieve only a small profit from implementing option B compared to that gained from implementing option A but now at the expense of a high comparative loss for her counterpart. Therefore, deception can be characterized as “mean” in this scenario. Lastly, in scenario 3 (high+,high-), the sender is able to gain a higher profit from option B than from option A at the cost of an equally high loss for her counterpart; therefore, in this case, a lie would be more economically justified than in scenario 2 even if it involves the same harm to the receiver.

Three different types of messages are identified. Because all messages state that a specific option will earn the receiver more money than the other two options, only one of them is actually an *honest* message. The remaining two messages can only be considered a lie. However, the message promoting the option which provides the highest payoff for the sender at the expense of a comparatively lower payoff for her counterpart involves a deceptive intention and will henceforth be called the *deceptive* message, whereas the message that promotes the Pareto-dominated option C is also a lie, but its main characteristic is to provide an equal payoff for both players and will therefore be labeled as the *payoff-equalizing* message.

3.3 Belief elicitation

Previous studies have shown that beliefs play an important role in understanding individual behavior in games associated with information asymmetry and misaligned payoff structures (e.g., Charness and Dufwenberg, 2006 & 2010, Sánchez-Pagés and Vorsatz, 2007, Hurkens and Kartik, 2009, Rode, 2010, López-Pérez and Spiegelman, 2013).

After making their decisions, but before presenting the final payoffs on the screen, both player types were asked to state their beliefs. First, we elicited subjects’ first-order beliefs regarding the behavior of their counterparts by asking senders to estimate the percentage of all receivers in their session who had accepted the messages and asking receivers about the percentage of senders who had sent truthful messages in their session.⁹ Second, we elicited the senders’ second-order beliefs about relative payoffs by asking how much they thought their counterpart expected to gain from the message relative to their own payoffs (providing five categories “much less”, “less”, “equal”, “more” and “much more”). In return, we asked receivers about their own relative payoff expectations to compare the senders’ second-order beliefs

⁸ The label indicates how much a sender earns (+) and a receiver loses (-) in purely monetary terms from an accepted deceptive message compared to an accepted honest message.

⁹ By eliciting beliefs regarding the share of trustful receivers and likewise the share of honest senders, we follow the approach of Rode (2010) and López-Pérez and Spiegelman (in press). This approach enables us to develop a subjective equilibrium analysis that we present in section 4.1.

with the actual receiver expectations. The importance of first- and second-order beliefs in a sender-receiver game analysis has recently been shown by Peeters et al. (2012).

Furthermore, we were interested in the subjects' peer group expectations because they may constitute a perceived "social norm". Therefore, after senders made their decisions about which message should be sent to their counterpart in each scenario, we showed them the three scenarios again and asked them on each of the screens to estimate (in four categories) how likely they believed it was that other players in their role had chosen a message that favored themselves in the respective scenario. In turn, receivers were asked about how likely they believed it was that other receivers in their session had accepted the message. The importance of peer group expectations has been shown, for example, by Lundquist et al. (2009) and Fosgaard et al. (in press) in sender-receiver settings as well as by Bicchieri and Xiao (2009) in a dictator game.

3.4 Treatments

The experiment consisted of three treatments. The general game structure and the payoff distribution for each scenario were used as a basis for all treatments. The only factor that changed among treatments was the amount of information that was revealed to the receiver after the game was played and, in particular, the probability with which this additional information was revealed. The first treatment (henceforth called T1) served as a control treatment in which no transparency about the payoffs was provided to the receiver except for her own final payoff at the end of the game.

In the second treatment (henceforth called T2), the game was played identically, except that now the receiver always found out about the honesty or dishonesty of the sender after the game was played. On the final screen of the game, we presented to the receiver her own final earnings and, additionally, all payoffs of both players in each of the options in the scenario that had been implemented. This subsequent disclosure was announced to both player types in the instructions.

In the third treatment (henceforth called T3), by default, we presented to the receiver only information about her own payoff corresponding to the implemented option at the end of the game, like in T1. Furthermore, there was a 50% probability that all information concerning the payoffs of both players in each option in the implemented scenario was revealed to the receiver at the end of the game, like in T2. All of the subjects knew about this probability of disclosure from the instructions.

3.5 Procedures

The experiment took place in 2011 at the Laboratorio de Economía Experimental (LEE), Universitat Jaume I, Castellón (Spain). Together, 468 undergraduate students were recruited using the Online Recruitment System for Economic Experiments ORSEE (Greiner, 2004). The subjects entered the laboratory one by one, sat down randomly in front of the computers and read the instructions. Nine sessions - three per treatment - were conducted, each with between 36 and 60 subjects (depending on the attendance each particular day). Our aim was to obtain at least 70 participants for each role for each treatment so as to obtain at least 70 independent observations per treatment and we obtained enough observations after the planned sessions. Depending on the treatment, the percentage of female subjects varied between 56% and 65% (see Table 2). The sessions lasted approximately 45 minutes. At the end of the session, subjects received their payoffs individually in cash. The average payoff was approximately 10€.

Treatments	Subjects	% of females
T1 - baseline Receivers never know the whole payoff structure	156	62%
T2 - subsequent disclosure Subsequent disclosure with 100% probability	144	56%
T3 - 50% subsequent disclosure Subsequent disclosure with 50% probability	168	61%

Table 2
Number of subjects and fractions of females per treatment.

4. THEORETICAL PREDICTIONS AND HYPOTHESES

In this section, we discuss strategies for each player type taking into account the sequential character and the information asymmetry of the sender-receiver game that we used in our experiment and derive the respective hypotheses.

4.1 Theoretical predictions for senders

In our design, the sender decides in each scenario $i \in \{1,2,3\}$ which of the available messages m_i she wants to send to the receiver. Each message m_i promotes one of the options z from the set of options $Z = \{A(\text{honest}), B(\text{deceptive}), C(\text{payoff} - \text{equalizing})\}$ of a particular scenario as the one that leads to the highest payoff for the receiver. However, the decision of whether this option will actually be implemented is up to the receiver. Hence, from the sender's perspective, there are mainly two sources of uncertainty: First, the unknown, subjective probability $p \in [0,1]$ with which the receiver accepts the message and, second, in the case of rejection, the probability with which the computer randomly selects one of the two remaining options - which we set to 50% in the experiment. These uncertainties are taken into account following the subjective equilibrium analysis of Kalai and Lehrer (1995) that was also used by Rode (2010) in a similar design. Equation (1) shows the sender's expected utility EU gained from sending message $m_i = z$ providing monetary payoffs $\pi_i(z)$ in scenario i under risk neutrality.¹⁰

$$EU(m_i = z) = p\pi_i(z) + (1-p)\frac{\sum_{j \neq z} \pi_i(j)}{2} \quad (1)$$

Figure 1 shows the sender's expected utility as a function of p according to equation (1) in scenarios 1 and 2. Taking into account only monetary payoffs, a sender who expects her counterpart to accept a message with p lower than 0.33 obtains the highest expected utility from sending the payoff-equalizing message. If a sender expects p to be higher than 0.33, the highest expected utility is achieved by sending a deceptive message. Senders with a probability expectation of exactly 0.33 are indifferent among the three messages. For all possible values of p , sending an honest message is never a dominant

¹⁰ The qualitative features of the model do not change when using constant absolute risk aversion or constant relative risk aversion utility functions.

strategy for a sender in any of the three scenarios.¹¹ Thus, according to standard economic theory, a sender should never send an honest message, except for the case of $p = 0.33$, in which she randomizes her decision.

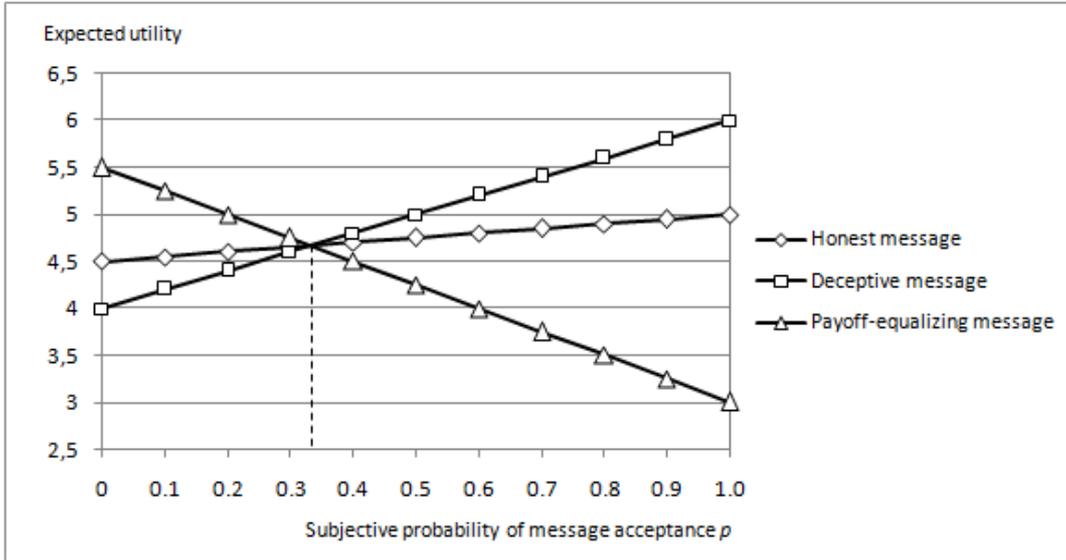


Figure 1
Sample simulation of the sender's expected utility according to subjective equilibrium analysis.

According to the lying cost approach (e.g., Kartik, 2009 and Lundquist et al., 2009), people who do not lie when facing a monetary incentive for doing so could suffer from a non-monetary cost that exceeds their expected payoffs from lying. Different factors driving these costs have been discussed in the literature, such as guilt aversion or aversion to lying *per se*. Some authors, such as Charness and Dufwenberg (2010) propose that a combination of both factors might have an effect on decision making. In addition to the purely monetary incentives, we assume that the decision to deceive is influenced by a general cost $c(i)$ that includes lying and guilt aversion as well as other intrinsic motivations. By letting $c(i)$ vary among different scenarios i , we also take into account the impact of the payoff distribution on senders' behavior. Given that the only design element changing among our treatments is the probability $d \in [0,1]$ with which *ex post* disclosure takes place, we assume that $c(i)$ affects all three treatments in the same way. We furthermore propose that the total costs of lying (C) also vary with social image costs $c_s(i;d)$ influenced both by the payoff distribution in scenario i and the probability of disclosure d .¹²

To capture all costs that affect the decision of a sender in our subjective equilibrium analysis, we introduce the total cost of lying function, depending on the message type $m_i = z$. Sending an honest message should not be negatively affected by any component of the lying cost function (equation 2). However, the sender's expected utility from sending deceptive messages is reduced by $C_B(c(i);c_s(i;d))$ (equation 3).

¹¹ The payoffs of the third scenario lead to a similar pattern with the same indifference point ($p = 0.33$).

¹² We introduce into the model a preference structure in which social image costs are different, although not necessarily independent, from other costs of lying. Of course, a different formulation in which transparency affects a more broadly defined concept of guilt aversion would also be possible.

In a similar way, we let the costs of lying $C_C(c(i);c_s(i;d))$ reduce the sender's expected utility from sending payoff-equalizing messages (equation 4).

Because sending deceptive and payoff-equalizing messages are both a lie but can have distinct characters in terms of their intention, the intensity of lying costs may vary between the two strategies. If the sender judges the deceptive message to be more provoking of guilt or lying aversion or social image concerns than the payoff-equalizing message, then $C_B > C_C$, whereas $C_B < C_C$ in the opposite case.

$$EU(m_i = A) = p\pi_i(A) + (1-p)\frac{\pi_i(B) + \pi_i(C)}{2} \quad (2)$$

$$EU(m_i = B) = p\pi_i(B) + (1-p)\frac{\pi_i(A) + \pi_i(C)}{2} - C_B(c(i);c_s(i;d)) \quad (3)$$

$$EU(m_i = C) = p\pi_i(C) + (1-p)\frac{\pi_i(A) + \pi_i(B)}{2} - C_C(c(i);c_s(i;d)) \quad (4)$$

In Figure 2, we represent the sender's expected utility by taking into consideration the costs of lying for a given $C_B > C_C > 0$. The expected utilities for both deceptive and payoff-equalizing messages are negatively affected by the costs of lying, and hence, sending an honest message becomes a dominant strategy in a specific area around the former indifference point $p = 0.33$, depending on the size of the lying costs arising in the other two options.

As an example simulation, in Figure 2, let us set $C_B = 0.3$ and $C_C = 0.15$ for scenario 1 as well as $C_B = 0.5$ and $C_C = 0.25$ for both scenarios 2 and 3. In this way, we account for the differences in payoff distributions among the scenarios and their effect on lying costs including $c(i)$ and $c_s(i;d)$. Because the receiver's monetary loss from successful deception is comparatively lower in scenario 1 than in the other two scenarios, we assume that the costs of lying are less affected in scenario 1. Correspondingly, we assume that the lying cost components are affected in the same way in scenarios 2 and 3 and hence control for the fact that the receiver's loss is equal between the two situations. Furthermore, the payoff that a sender can gain from lying is the same in scenarios 1 and 2, and therefore, the aforementioned lying cost effect leads to a comparatively wider p -range in which honesty becomes a dominant strategy (between the dashed lines in Figure 2) in scenario 2 compared to scenario 1. In scenario 3, senders are able to gain a comparatively higher profit from sending a deceptive message. As a consequence, the value range p for which honesty becomes a dominant strategy is smaller in scenario 3 than in the other two scenarios.

This simulation of the subjective equilibrium analysis is in line with previous findings indicating that people do not only care about their own monetary payoffs but also about how they can achieve them. The different sizes of the p -ranges in which honesty is a dominant strategy for senders capture the patterns of honesty among similar payoff scenarios observed by Gneezy (2005).

The lying cost component representing social image concerns $c_s(i;d)$ also depends on the probability with which conflicts of interest and the sender's actual behavior are revealed. In our subjective equilibrium

analysis, we let values of $d > 0$ cause a lower expected utility from sending deceptive and payoff-equalizing messages for subjects who are sensitive to a potential loss of their social reputation in this context. Following the patterns among the different payoff scenarios that we developed previously, this situation leads to correspondingly larger p -ranges in which honesty becomes a dominant strategy (as shown in Figure 2). A priori it is reasonable to assume that this effect becomes stronger with an increasing probability d because subjects who suffer from a cost of lying driven by social image concerns should be less inclined to lie if the probability of a disclosure of their behavior increases.

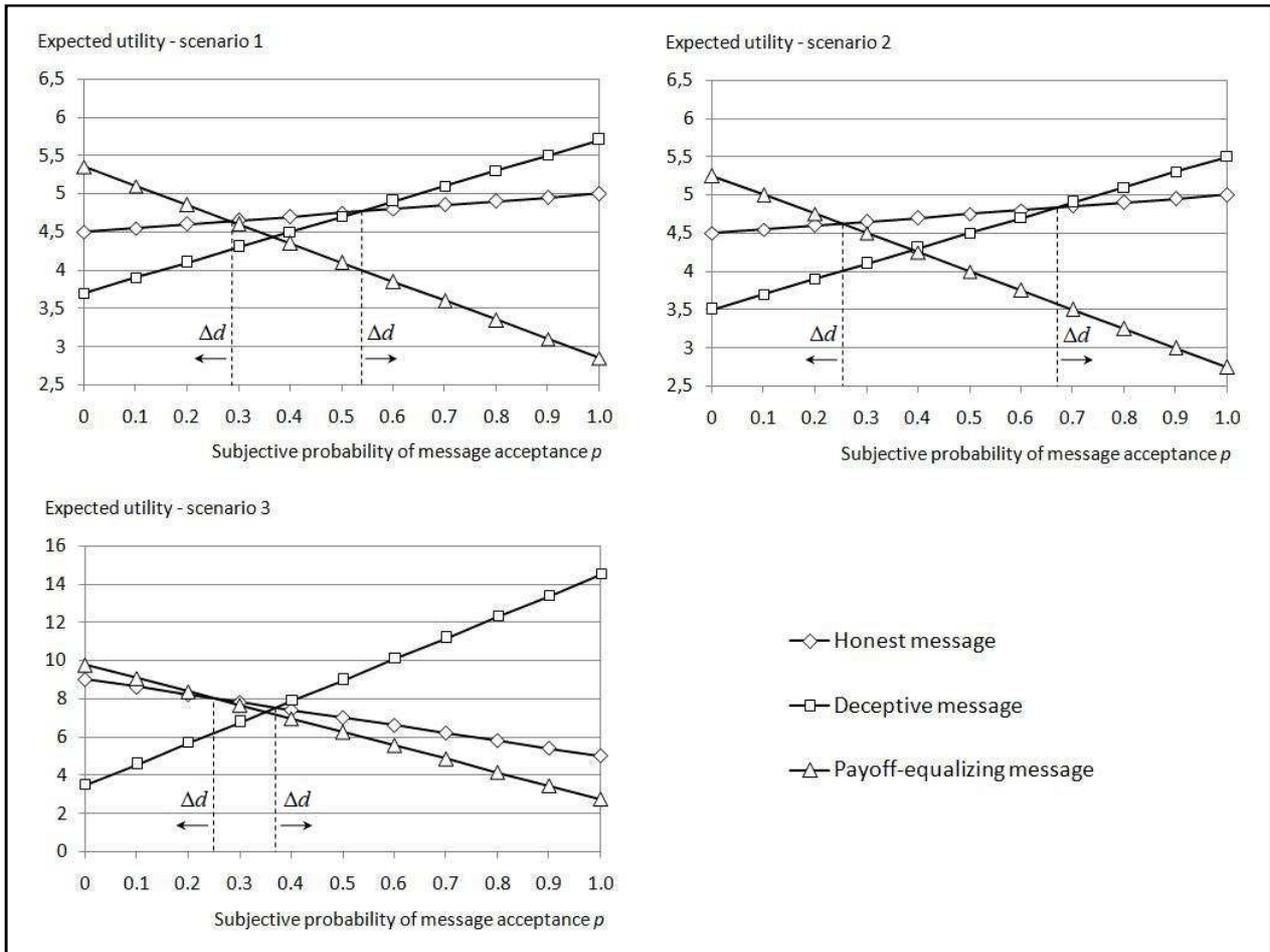


Figure 2
Sample simulations of the sender's expected utility according to subjective equilibrium analysis including lying costs.

4.2 Theoretical predictions for receivers

In contrast with other designs such as that reported in Rode (2010), in which receivers had some information about their possible payoffs in several treatments, our receivers are blind regarding payoffs. Therefore, their decision depends only on their beliefs regarding the honesty of the received message. In this sense, no specific expected payoff calculation is applicable.

Let $q \in [0,1]$ be the probability with which the sender transmits an honest message. If the receiver thinks that q is greater than 0.5, i.e., her counterpart is more likely to have sent an honest message than a lie, she maximizes her expected payoffs by accepting the message and vice versa. For $q = 0.5$, she should be indifferent between accepting and rejecting a message.

4.3 Testable hypotheses

The subjective equilibrium analysis of the expected payoffs indicates that sending honest messages is never the dominant strategy in our setting if we only take into account monetary incentives. However, in our extended model with costs of lying, we show that sending an honest message can become a dominant strategy for some senders given particular beliefs and depending on the scenario. According to this model, we expect to find similar patterns in our experiment. In particular, although it is reasonable to assume that social image concerns will not play a significant role in the baseline treatment, we expect this element to increase the total costs of lying in our second and third treatments in which we provide *ex post* transparency with specific probabilities.

Hypothesis 1a:

Compared to non-disclosure, subsequent disclosure leads to a reduction in the rate of deceptive messages.

In line with our theoretical predictions, we assume that deception decreases with an increasing probability of subsequent disclosure d because the social image effect should, in principle, become stronger with an increase in the probability that the sender's behavior will be revealed to the receiver.

Hypothesis 1b:

The reduction in deception is higher in the case of assured disclosure (T2) than in the case with 50% probability of disclosure (T3).

Furthermore, social image concerns could increase the impact of deception further when the consequences of the lie are worse for the receiver, as it is the case in scenarios 2 and 3.¹³ However, the effect in scenario 3 will probably be counterbalanced by the possibility that the sender obtains much higher gains for herself by lying in that particular scenario.

Hypothesis 1c:

The increase in the rate of honest messages with subsequent disclosure is expected to be greater in scenario 2 (low+;high-) than in the other scenarios.

Furthermore, given our previous theoretical analysis, we expect that sending payoff-equalizing messages is the dominant strategy for low perceived probabilities of the receiver accepting the message, that truth telling is dominant for intermediate probabilities and that sending deceptive messages is dominant for high probabilities.

Hypothesis 1d:

It is more likely that a sender transmits a payoff-equalizing message for low than for high first-order beliefs in all scenarios, and it is more likely that a sender transmits an honest message for intermediate than for high first-order beliefs in all scenarios.

Turning to receiver's behavior, we expect that a receiver will anticipate the potentially positive effect of subsequent disclosure in terms of reducing deception in T2 and T3.

¹³ This effect could also be interpreted as an interaction of social image concerns with guilt aversion.

Hypothesis 2a:

Compared to that in the non-disclosure case, the message acceptance rate increases with subsequent disclosure.

Given that receivers have no *ex ante* information about payoffs, their first-order beliefs about the probability that senders sent an honest message are expected to be the main driving force behind the decision whether to accept the message.

Hypothesis 2b:

The higher their first-order beliefs, the more likely it is that receivers accept the message.

5. RESULTS AND DISCUSSION

In this section, we first conduct a descriptive analysis of the sender's behavior in each scenario and per treatment. Second, using the individual data, we estimate an econometric regression based on our theoretical model, which validates our results and provides further insight into the moderating role of individual beliefs in this context. We lastly provide an analysis of the receiver's behavior and beliefs in our experiment.¹⁴

5.1 Behavior of senders

Figure 3 illustrates the percentages of honest, deceptive and payoff-equalizing messages sent in each scenario per treatment. We start by analyzing the fractions of deceptive messages. In general, we do not find any increase in the rate of deceptive messages in the treatments with *ex post* transparency compared to the baseline. Therefore, the so-called "license to lie" effect of *ex ante* transparency, which has been observed by many other authors (e.g., Cain et al., 2005), does not appear in our design with subsequent disclosure as shown in Figure 3.

From the between-subject perspective, i.e., how different degrees of subsequent disclosure affect subjects' decisions in a specific scenario, we find almost no differences in the rates of deceptive messages in the first scenario. In all treatments, approximately 50% of the subjects chose deceptive messages in this scenario, in which subjects play for low stakes. In scenario 2, we find a significant reduction of 13% in deceptive messages, from 55% in T1 to 42% in T3, in which subsequent disclosure is implemented with 50% probability (a chi²-test shows a p-value of 0.087). A similar pattern can be observed in scenario 3: Compared to the baseline, the rate of deception decreases in T2 and T3, however, this decrease is not significant. A possible explanation is that the high gain from lying that can be obtained by the sender in scenario 3 counteracts the social image effect of subsequent disclosure, making the observed decrease in the rate of deceptive messages insufficiently substantial to be significant. In light of these results, we confirm our Hypothesis 1a partially because the hypothesized reduction in lying occurs only in one particular scenario and treatment. Because we find no lower rates of deceptive messages in T2 than in T3, we have to reject our Hypothesis 1b.

¹⁴ The raw data and a list of the variables collected are available at www.lee.uji.es/eng/personales/ivan/subsequent_disclosure.xlsx. We did not eliminate any data obtained for performing our analysis.

Result 1:

Compared to T1, the rate of deceptive messages never increases and it significantly decreases only in T3 and scenario 2.

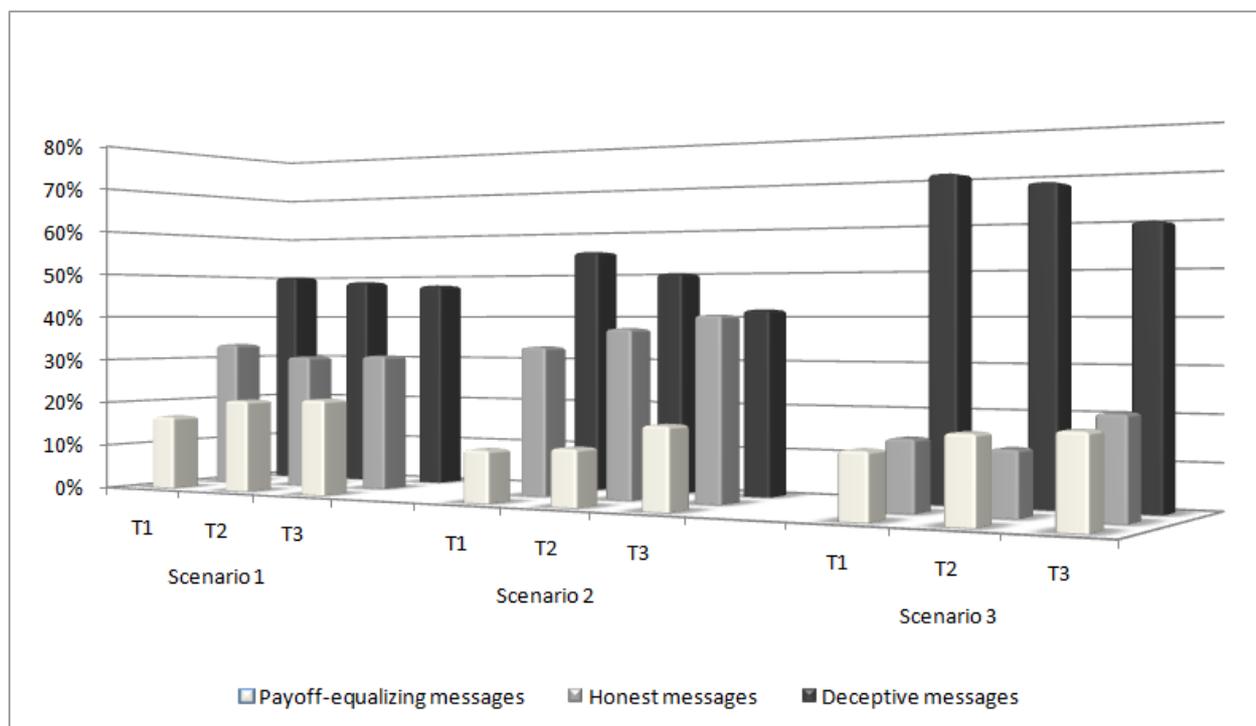


Figure 3
Fractions of messages sent per scenario and treatment.

What this result tells us is that subsequent disclosure does not lead to an increase in deception and that it reduces deception only under two simultaneous conditions: (i) when the counterpart might be greatly harmed simply for a relatively small gain from deception for the sender and (ii) when subsequent disclosure is implemented with 50% probability. A reasonable explanation for the first condition stems from the specific characterization of deception in this payoff scenario, which can be perceived as “mean”, taking into account the disproportion between gains for the sender and losses for the receiver from a lie in this situation. This perception might create or intensify the social image effect for some of the subjects. The second condition for a significant reduction in deception by *ex post* transparency is more puzzling. Taking into account the findings reported in the literature about perception of probabilities and risk, a potential explanation for this result could be that mentioning a specific numerical probability of transparency instead of an assured disclosure creates more awareness of the implications of deception and hence has a stronger effect on reducing it than disclosure with 100% probability, a situation in which the participants are not necessarily aware of the counterfactual.¹⁵ Another possible explanation is that the perspective of being one among all people who have lied in the treatment, and who then become exposed by assured subsequent disclosure, could produce less anticipated social image concerns than envisaging being one of the relatively few liars out of the total who are actually discovered by a 50% transparency rate. In fact, out of 72 and 84 sender participants in the treatments with subsequent transparency T2 and T3, the 41

¹⁵ For example, Fiedler (1988) finds that a modification of subtle linguistic factors can substantially reduce biases in understanding probability concepts, in this case regarding the conjunction fallacy. Stone et al. (1997) examine how different framings of low-risk probabilities affect individual behavior. Weber et al. (2005) show that the perception of investment risk differs significantly depending on type and presentation of the given information.

liars were caught in T2 and only 17 out of 40 liars were uncovered in T3. More research is needed to explore in detail the sensitivity of the costs of lying to the probability of disclosure. At this point, with just our three treatments, we are not able to establish whether the relationship is linear or even monotonic.

Another interesting outcome is that we find significant gender differences in scenario 2. In T2, 54% of the female subjects sent deceptive messages, whereas only 39% of the male senders did so. In T3, we find a similar pattern because 51% of the female subjects and only 27% of male senders decided to send deceptive messages. It seems as if the male subjects are more sensitive to a “mean” deception when *ex post* transparency is provided.¹⁶

With regard to honesty, we find a non-negligible fraction of subjects who sent truthful messages in all scenarios and treatments (between 14-40%), which indicates that a substantial part of the subjects experience positive costs of lying. Although not significant, Figure 3 shows that an increase in the rate of honest messages with subsequent disclosure compared to the baseline appears in scenario 2 and only for T3, whereas a higher fraction of payoff-equalizing messages compared to the baseline is found in all three scenarios and in both disclosure treatments. This result could be caused by the 50-50 rule (see Bénabou and Tirole, 2006) being perceived as less damaging to social image than lying, and at the same time, the equalitarian message avoids the envy experienced when recommending the best option to the principal.

When testing for within-subject effects of different payoff misalignments on deception, i.e., comparing behavior among scenarios within a treatment, we find a significantly higher percentage of dishonest messages in scenario 3 than in scenario 1 and 2 for all treatments.¹⁷

Result 2:

In scenario 3, in which the sender gains a comparatively high profit from deception, the rate of deceptive messages is significantly higher than in the other two scenarios - regardless of the implementation of subsequent disclosure.

A similar result was shown by Gneezy (2005): “This unselfish motive diminishes with the size of the gains to the decision maker herself”.¹⁸

The fractions of honest messages are significantly higher in the first two scenarios than in scenario 3, in which the sender gains a high amount from successfully deceiving the receiver.¹⁹ In contrast, the differences among the rates of payoff-equalizing messages across scenarios are not significant except for the difference between scenario 1 and 2 in T2 ($p=0.058$).

Result 3:

The tendency to send honest messages is more influenced by the scenarios’ different payoff distributions than the tendency to send payoff-equalizing messages.

¹⁶ The p-values of the respective χ^2 -tests are 0.098 for T2 and 0.031 for T3. Dreber and Johannesson (2008) find that more male than female subjects lie in a sender-receiver game with payoffs comparable to those of our scenario 1. A similar result regarding “selfish black lies” was also shown by Gneezy and Erat (2012). On the other hand, the results of Childs (2012) do not show gender differences in a payoff setting comparable to that of our scenario 3. Note that there was no third payoff-equalizing option used in these studies.

¹⁷ A McNemar test shows p-values of 0.002 for T1, 0.003 for T2 and 0.041 for T3 when comparing scenario 1 and 3. The respective p-values for a comparison between scenario 2 and 3 are 0.014 for T1, 0.016 for T2 and 0.005 for T3.

¹⁸ Compared to Gneezy (2005), our results show higher rates of deception in all scenarios. One difference with our design is that Gneezy did not use the Pareto-dominated third option in his scenarios, but this should rather decrease than increase deception. Other distinctions from Gneezy’s study are possible cultural differences regarding the subjects as well as the fact that Gneezy uses a between-subject design to compare different scenarios, whereas we use a within-subject design.

¹⁹ The p-values of the McNemar test for the comparison of honest message fractions between scenario 1 and scenario 3 (scenario 2 and scenario 3) are 0.011 (0.006) in T1, 0.131 (0.003) in T2 and 0.014 (0.001) in T3.

5.2 The role of senders' beliefs

In the following paragraphs, we analyze (i) the sender's beliefs about the behavior of receivers, (ii) their beliefs about their counterpart's relative payoff expectations, i.e., how much they think the receiver expects to gain from the message in relation to the sender's profit, and (iii) their expectations about peer group behavior, i.e., how likely they think it is that other senders would send a deceptive message in a given scenario. Table 3 summarizes the means of the first-order beliefs and the percentages of other belief categories across treatments. For all categories of the three belief types, none of the differences among treatments is significant at conventional levels. Hence, the observed reduction of deceptive messages with 50% *ex post* transparency in scenario 2 cannot be explained exclusively based on changes in beliefs across treatments.

Result 4:

Concerning senders' beliefs, no significant differences across treatments are observed.

Beliefs of senders	Treatments		
	T1	T2	T3
First-order beliefs about the receiver's action	Means		
Percentage of receivers that accept the message	44.91	45.18	41.88
Second-order beliefs about relative payoff expectations	Percentages		
<i>Lower or much lower</i> than sender's payoffs	53.85	55.56	58.33
<i>Equal</i> to sender's payoffs	16.67	15.28	19.05
<i>Higher or much higher</i> than sender's payoffs	29.49	29.17	22.62
Peer group beliefs	Percentages		
Other senders <i>likely/very likely</i> to send deceptive messages			
Scenario 1 (low+;low-)	75.65	76.39	76.19
Scenario 2 (low+;high-)	78.20	68.05	73.81
Scenario 3 (high+;high-)	69.23	68.06	71.42

Table 3

Means and percentages of senders' beliefs across treatments.

Nevertheless, beliefs can still be a driving force for the sender's behavior at an individual level. To analyze these possible influences, we now turn to the subjective equilibrium analysis of the sender's behavior developed in section 4.1, which is mainly based on the sender's first-order beliefs. According to these predictions, senders should never send honest messages intentionally except when they suffer from a psychological cost of lying. In the extended model with costs of lying, honesty becomes a dominant strategy for intermediate probabilities p with which receivers are expected to accept the message - depending on their individual costs of lying -, whereas sending payoff-equalizing (deceptive) messages continues to be the dominant strategy for low (high) values of p .

Figure 4 presents the distribution over treatments and scenarios of the number of senders who chose the respective message type per quintile of the first-order beliefs p (in percentage). It can be observed that deceptive messages concentrate and dominate particularly in intermediate and high quintiles of the p -distribution, whereas payoff-equalizing messages reach their highest proportions in the first and second quintile in almost all scenarios and treatments. Honest messages tend to concentrate in intermediate quintiles of the p -distribution. In scenario 2, where we find that the implementation of *ex post* transparency with 50% probability significantly reduces deception, honest messages are indeed the most frequently chosen messages in the second quintile of the distribution both in T2 and in T3. These observations are in

line with our extended model and Hypotheses 1c and 1d. This finding implies that some senders suffer from an additional cost of lying caused by subsequent disclosure through social image effects, particularly when they can inflict their counterpart great harm for little gain, and hence choose to send an honest message.

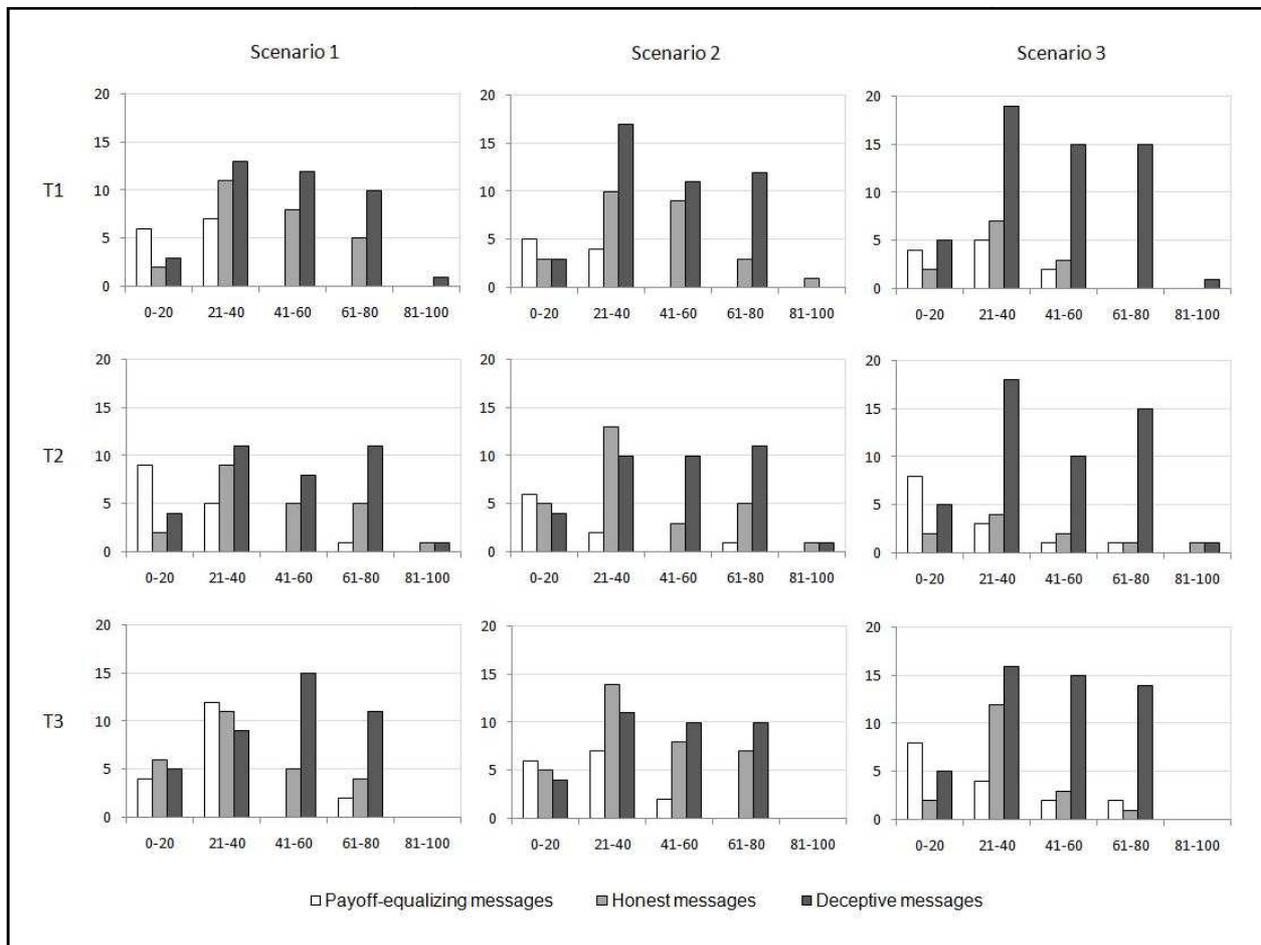


Figure 4 Distribution of message types per quintile of the p -distribution in each scenario and treatment (in percentages).

We now turn to a more encompassing econometric model which takes into account for each particular scenario the main factors that could be interacting to affect the sender's decision and for which we have obtained data: subsequent transparency (captured by treatment dummies), first-order, second-order and peer beliefs as well as socio-demographic variables.

We use multinomial logistic regressions to control for the polytomous character of the dependent variable, which is the sender's behavior in terms of the chosen message.²⁰ The model enables us to choose deception as a base category for examining the likelihood that subjects will switch to the alternative of sending honest messages or payoff-equalizing messages for different levels of the independent variables. The results are shown in Table 4. The coefficients show the change in the relative log-odds between the base outcome and the categories honesty and payoff equalization.²¹

²⁰ There is no unique order among the values of the dependent variable because it depends on individual preferences which of the lies is considered worse: harming a counterpart to receive a higher payoff or reducing the payoffs for both players equally to a substantially lower level.

²¹ The first-order belief variables correspond to the quintiles of the distribution of p , the sender's subjective probability with which the receiver accepts the message, which we used in our theoretical framework. Quintiles Q4 and Q5 of the distribution have been

Treatment and belief effects		Scenario 1 (low+;low-)	Scenario 2 (low+;high-)	Scenario 3 (high+;high-)
Honesty	T2	-0.061 (0.382)	0.330 (0.373)	-0.292 (0.481)
	T3	-0.038 (0.375)	0.740** (0.375)	0.374 (0.432)
	First-order_quintile 1	1.486* (0.774)	0.772 (0.768)	2.301*** (0.833)
	First-order_quintile 2	0.573 (0.387)	0.854** (0.369)	1.337*** (0.460)
	First-order_quintile 3	0.528 (0.401)	1.121*** (0.415)	0.944* (0.523)
	Second-order_more	1.052*** (0.347)	0.0465 (0.342)	0.619 (0.417)
	Peer_group_lying	-0.824** (0.324)	-1.401*** (0.321)	-2.067*** (0.433)
	Female	0.048 (0.331)	-0.492 (0.321)	0.142 (0.387)
	Siblings_max one	-0.522 (0.353)	-0.169 (0.347)	-0.558 (0.414)
	Economics_Business	-0.621* (0.372)	0.163 (0.339)	-0.820* (0.451)
Grant	0.104 (0.355)	-0.060 (0.346)	-0.190 (0.436)	
Deception		- base outcome -		
Payoff-equalization	T2	-0.494 (0.471)	-0.397 (0.547)	-0.098 (0.509)
	T3	-0.217 (0.454)	0.314 (0.516)	0.218 (0.487)
	First-order_quintile 1	2.869*** (0.789)	3.248*** (0.795)	3.788*** (0.800)
	First-order_quintile 2	1.968*** (0.459)	1.979*** (0.539)	1.884*** (0.533)
	First-order_quintile 3	0.107 (0.621)	1.070 (0.682)	1.589** (0.618)
	Second-order_more	0.366 (0.464)	-0.096 (0.497)	0.788* (0.461)
	Peer_group_lying	-1.509*** (0.380)	-1.962*** (0.439)	-2.809*** (0.473)
	Female	-0.237 (0.391)	-0.853* (0.454)	-0.520 (0.422)
	Siblings_max one	-1.110*** (0.415)	-0.758* (0.452)	-0.242 (0.461)
	Economics_Business	0.332 (0.411)	-0.382 (0.506)	-0.737 (0.468)
Grant	-0.205 (0.461)	-1.164* (0.627)	-0.462 (0.509)	
Wald	68.62***	75.26***	109.66***	
N	234	234	234	

Note: *** p-value < 0.01; ** p-value < 0.05; * p-value < 0.1. Standard errors in parentheses.

Table 4

Multinomial logistic regression model for senders' behavior and beliefs.

Treatment effects

The coefficient for the treatment dummy variable T3 shows that the provision of *ex post* transparency with 50% probability, after controlling for the possible effects of changes in individual beliefs, is to significantly increase the relative probability that subjects change from deception to honesty in scenario 2, where deception can be characterized as “mean”. This econometric result further confirms our Hypotheses 1a under those conditions. Moreover, we reject our Hypothesis 1b and confirm our Hypothesis 1c. *Ex post* transparency does not have an effect on the probability of sending a payoff-equalizing message further than that already captured in the changes in beliefs.

First-order beliefs

Table 4 shows highly significant positive effects of first-order beliefs on the probability of sending payoff-equalizing messages for low values of p in all scenarios. In scenario 3, this is also true but less significant for intermediate values of p . With regard to honest messages, senders who believe that receivers will

combined due to few observations in Q5. They act as a baseline for comparison. The variable *second-order_more* acts as a dummy variable for senders who believe that their counterparts expects to achieve a higher payoff from the message compared to the sender' gain. *Peer_group_lying* acts as a dummy variable for senders who expect the majority of their peers to lie in a specific scenario. The dummy *siblings_max one* captures senders who have more than one sibling. Socio-demographic variables which resulted insignificant are not presented in the table but the results are qualitatively similar in their presence.

accept messages with a low or intermediate value of p are significantly more likely to send honest messages than those with beliefs in higher quintiles. These findings are broadly consistent with our Hypothesis 1d.

Result 5:

Subjects with low values of first-order beliefs are more likely to send payoff-equalizing messages than subjects with high values. Furthermore, subjects with low or intermediate first-order beliefs are more likely to send honest messages than subjects with high values.

Second-order beliefs

In general, beliefs about the counterpart's relative payoff expectations do not have a substantial effect on sender's behavior. Table 4 shows that the only significant effect of second-order beliefs on honest behavior can be found in scenario 1 in the sense that if senders believe that their counterparts expect to achieve a comparatively higher payoff (variable second-order *more*), the relative probability of being honest increases. Regarding payoff-equalizing messages, we only find a slightly significant effect of second-order beliefs in scenario 3. We hypothesize that this low relevance of second order beliefs is a reflection of the senders knowing that in our experimental design receivers did not have any information about the possible payoffs at the beginning of the game.

Peer group beliefs

An important factor that increases the relative probability that subjects change from deception to honesty is the sender's expectation about the action of her peers, which could have a direct effect on lying, guilt and social image concerns. We assimilate the senders' conformity to the expected actions of their peers with their particular sense of accordance to a perceived social norm in a given scenario and treatment. In the regression above, we obtain that higher expectations about the probability that other players in the same role will send deceptive messages (variable *peer_group_lying*) have a highly significant negative effect on the likelihood of sending honest and, respectively, payoff-equalizing messages.

Result 6:

The more a sender believes that the other players in the same role will send deceptive messages in a specific scenario and treatment, the higher the probability becomes that she will send a deceptive message as well.

This result provides an interesting hint indicating that social norms in a group play a role when deciding whether to lie. It seems that players tend to be reluctant to deceiving when they actually believe that they are breaking a legitimized social norm when they deceive. Lundquist et al. (2009) find a similar effect on lying aversion through the acceptance of social norms in peer groups.

Socio-demographic variables

When taking into account the role of beliefs, gender only affects sending payoff-equalizing messages in scenario 2. Female subjects in scenario 2 are less likely to promote the Pareto-dominated option. In scenarios 1 and 2, in which senders gain a comparatively low profit from successful deception, we find that the probability of sending a payoff-equalizing message decreases when the sender has at most one sibling. The subject's academic background affects honesty in the sense that if subjects come from a economics and business background, they are less inclined to send honest messages when the loss for

the receiver is low (scenario 1) and when stakes are high (scenario 3). This result that students with economics background tell the truth significantly less often is in line with López-Pérez and Spiegelman (2012) albeit not for the mean scenario 2 in our study. Furthermore, we find a significant negatively effect of being a grant-holder only for payoff-equalizing messages in scenario 2.²²

5.4 Receiver behavior

In general, we observe low acceptance rates compared to other studies concerning similar sender-receiver games.²³ As shown in Figure 5, 42% of the receivers accepted the messages sent by their counterparts in T1. In T2, the percentage of acceptance decreased to 33%. This is a surprising result, because we expected that the probable deception reducing effect of subsequent disclosure in T2 would have been anticipated by the receivers. Nevertheless, it could be the case that some receivers would like to avoid the possibility of finding out that they have been successfully deceived, and in the case of assured disclosure, this can only be achieved by rejecting the message. This possible explanation is consistent with our observation in T3, in which the acceptance rate of 48% is higher than that in the other two treatments, as if receivers had correctly anticipated that the reduction in dishonest messaging would be greater in T3 than in T2 and the probability of finding out that you have been cheated upon is lower or equal than 50%.²⁴ However, the differences between the baseline and the treatments with disclosure are not significant at conventional levels. Therefore, our data do not support Hypothesis 2a.

Result 7:

Compared to the baseline, the percentage of receivers accepting the message with subsequent disclosure is not significantly different.

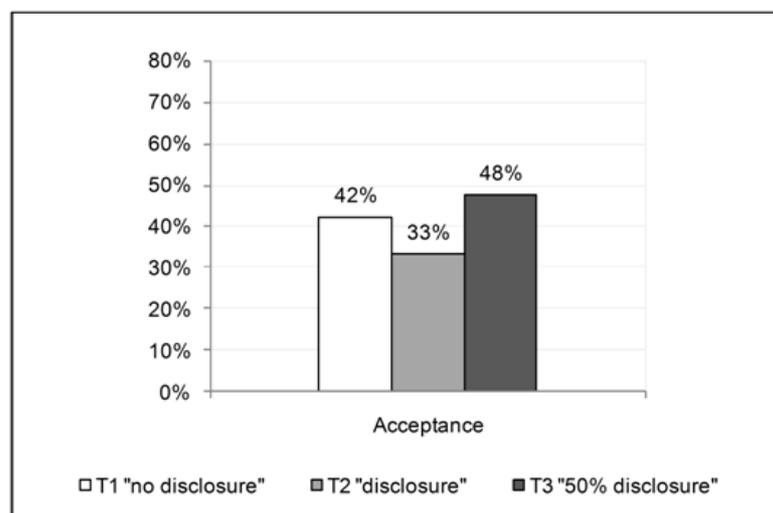


Figure 5

Acceptance rates of receivers per treatment.

We observe interesting gender differences in acceptance rates between the treatments. Compared to only marginal differences in acceptance rates for female subjects in all treatments (between 39% and 41%), we

²² The questionnaire that all subjects had to complete during the sessions also included also questions regarding lying in general and emotions experienced under different hypothetical scenarios. These subjective-non incentivized data did not lead to any significant result during the data analysis.

²³ For example, Gneezy (2005) finds that 78% of the receivers follow the advice of the senders. In Dreber and Johannesson (2008), 76% of the receivers accepted the message. On the other hand, Inderst et al. (2010) also observe lower acceptance rates between 47% and 51%, depending on the degree of *ex ante* disclosure.

²⁴ A chi²-test shows a p-value of 0.071 for the difference between T2 and T3.

find that the acceptance rate for male subjects increases significantly from 26% in T2 to 55% in T3.²⁵ Compared to the case of assured subsequent disclosure, male receivers correctly anticipate the effectiveness of subsequent disclosure with 50% probability in reducing deception. This finding is in line with the behavior of male senders, who also seem to be more sensitive to subsequent disclosure in T3.

5.5 The role of receiver's beliefs

Table 5 presents the elicited beliefs of subjects in the role of receivers. According to the first-order beliefs, only one-third of the receivers thought that their counterparts sent an honest message in the first two treatments (41% in T3). Furthermore, over 40% of the receivers expected to gain a comparatively lower payoff from the sent message. Regarding peer group beliefs, we find that between 43% and 46% of the receivers expected the other players in their role to accept the messages they received. As in the case of senders, the receivers' beliefs do not significantly differ among treatments.

Beliefs of receivers	Treatments		
	T1	T2	T3
Counterpart beliefs	Means		
Percentage of senders that sent an honest message	36.82	34.50	40.46
Relative payoff beliefs	Percentages		
<i>Lower or much lower</i> than receiver's payoffs	43.59	41.67	42.86
<i>Equal</i> to receiver's payoffs	24.36	29.17	28.57
<i>Higher or much higher</i> than receivers' payoffs	32.05	29.17	28.57
Peer group beliefs	Percentages		
Receivers <i>likely/very likely</i> to accept sent messages	46.15	43.06	42.86

Table 5
Means and percentages of receivers' beliefs across treatments.

As in the analysis of the senders' behavior and beliefs, we now estimate an econometric model that captures the effect of the transparency treatments, first-order, relative payoff and peer beliefs as well as demographic variables.²⁶ The results of this logit model are shown in Table 6. According to the estimated coefficients, the provision of subsequent disclosure does not affect the likelihood of the receiver accepting the message. Furthermore, Table 6 shows that the receivers' first-order beliefs have a highly significant effect on acceptance in the sense that the higher receivers rate the probability that senders will transmit an honest message, the more probable it is that they accept the received message. This result confirms our Hypothesis 2b.

Result 8:

The higher the receiver's first-order beliefs, the more likely it is that she will accept the message.

With respect to relative payoff expectations, we find a significant effect on acceptance rates in the sense that the probability of acceptance decreases if the receiver expects to gain a higher amount from the message compared to the sender's payoff. This result is in line with the comparatively low acceptance rate that we observe in the sense that selfish receivers seem to distrust their counterparts. Other than in the case of the senders, peer beliefs do not explain the behavior of receivers in our experiment.

²⁵ A chi²-test shows a p-value of 0.019. Our results do not show significant differences between the acceptance rates of male and female subjects within treatments and hence generally confirm the findings of Dreber and Johannesson (2008) and Childs (2012).

²⁶ Dependent variable: message acceptance. N indicates the number of receivers.

Treatment and belief effects on message acceptance		
T2	-0.321	(0.420)
T3	0.063	(0.408)
First-order_beliefs	0.069***	(0.010)
Relative payoff_more	-0.789*	(0.408)
Peer_group_lying	0.185	(0.353)
Female	-0.431	(0.367)
Siblings_max one	-0.069	(0.369)
Economics_Business	-0.424	(0.379)
Grant	-0.072	(0.458)
Constant	-2.354***	(0.594)
Wald	96.64***	
Pseudo R ²	0.3044	
N	234	

Note: *** p-value < 0.01; ** p-value < 0.05; * p-value < 0.1.
Standard errors in parentheses.

Table 6
Logistic regression model for the receiver's behavior and beliefs.

6. CONCLUSIONS

Asymmetric information often provides incentives to deceive in the information transmission process between principals and their agents. Our main interest in this study is to investigate how these incentives may be counteracted by the provision of transparency. Because previous results reported in the literature show that *ex ante* transparency about the conflict of interest may even lead to a higher rate of deception, we propose a theoretical framework and conduct an experiment to examine the effect of a new modification in a sender-receiver game intended to reduce treacherous behavior: *ex post* transparency.

Our results show that, although all subjects remain anonymous, the rate of deceptive messages does not increase when the payoff structure of the game is disclosed to receivers after the interactions. The so-called "license to lie" effect, which has been observed in studies concerning *ex ante* transparency, does not appear with subsequent disclosure, which can, in principle, therefore serve as a policy instrument for further measures, such as deterrence or as a basis for penalties. In particular, *ex post* transparency even reduces deception significantly: in the case in which it is implemented with 50% probability and deception can be characterized as "mean", i.e., when a small incremental profit (1€) can be gained by the sender at the expense of a relatively greater harm (10€) caused to the payoff of her counterpart. The character of the payoff misalignment also provides an explanation for the fact that we do not observe a significant decrease in deception in the other scenarios. Because the difference between the low gain of the sender (1€) and the equally low receiver's loss (1€) in scenario 1 is small, the consequences for the receiver very often do not serve as an effective incentive to keep senders from deceiving (with or without transparency). When stakes are as high as in scenario 3, in which 10€ are transferred from the receiver to the sender by a successful lie, subjects are also more inclined to deceive their counterparts to a point that *ex post* transparency cannot remediate. Economics and business students are slightly less prone to send honest messages than students with other background in the cases when the harm to the receiver is small or the gain they can make is great.

We propose that an intrinsic motivation that goes beyond lying and guilt aversion plays a role when implementing *ex post* transparency about conflict of interests: social image concerns.

In classical sender-receiver games used in past experiments, receivers did not find out about conflicts of interests after the game and hence did not know whether the sender was in fact honest. In those situations, the sender had no reason to believe that her social image could suffer from acting immorally because the receiver remained blind regarding the ethical character of her decision. Senders, who sent honest messages in these experiments, rather seem to be driven by motivations such as lying aversion that are not necessarily connected to their social image.

However, in the line of experiments in which receivers are informed about conflicts of interest before the game, the fraction of dishonest senders increases compared to that in other situations, in which no such *ex ante* information was revealed to the receiver. As a possible explanation for this robust pattern, the authors refer to the effect of moral licensing. "With disclosure of a conflict of interest, giving biased advice might seem like fair play" (Cain et al., 2005). If the otherwise unethical action is believed to be perceived as fair by the receivers, senders should not be concerned about their social image.

Because receivers are not provided with pre-play information about the conflict of interest in this study and hence no moral licensing is possible, it is reasonable to assume that *ex post* disclosure of the actual behavior of the sender affects her social image in this case. Therefore, the audience effect through the receiver as a directly affected observer is a promising explanation for our experimental findings.

As a tentative explanation for the increased effect of 50% transparency, we propose that simply the act of mentioning the two alternative worlds, with and without transparency, makes participants think more about the implications of deception than in the other treatments in which only one of the two alternatives is presented to them. Another possible explanation is that the perspective of being one among all people who have lied in the treatment, and who then become exposed by assured subsequent disclosure, could produce less anticipated social image concerns than envisaging being one of the relatively few liars out of the total who are actually discovered by a 50% transparency rate. Further research varying the transparency rate is needed in order to shed light on the particular form of the relationship between deception and the probability of ex-post disclosure.

To develop a well-controlled explanation of why subsequent disclosure keeps people from deception, we also take into account the moderating role of individual beliefs in an econometric model. We find that the rate of honest messages, rather than that of payoff-equalizing ones, increases when deception is significantly reduced by subsequent disclosure. In line with our theoretical predictions, we find that the subjects' likelihood to send honest and payoff-equalizing messages is greater for low and intermediate first-order beliefs, whereas deceptive messages are preferred for high first-order beliefs. We also find that beliefs about peer actions have a strong and significant effect on the likelihood of sending deceptive messages in all scenarios in our experiment: the more a sender believes that other players in her own role will deceive, the higher the probability that she sends a deceptive message herself. These results are in line with the implications of other studies. For instance, Bicchieri (2006) states that guilt not only depends on violating a social norm but also on this norm being considered legitimate. In addition, Gneezy and Erat (2012) state, "People feel guilty when they lie and violate a social norm. The amount of guilt they feel depends on the descriptive norm, i.e., their beliefs about the adherence to the norm in their peer group." In

this sense, some people would like to avoid situations in which their deceptive behavior is disclosed after breaking a social norm that they think is widespread among their peers.

With regard to the behavior of receivers, we find that they do not anticipate the decrease in deception that subsequent disclosure provides. In fact, assured *ex post* transparency leads to a substantial but not significantly lower acceptance rate than in the case of no transparency, as if receivers want to avoid finding out that they have become a victim of deception. We find that first-order beliefs significantly affect the behavior of receivers in the expected way: the higher the probability assigned by receivers to senders having sent an honest message, the more likely it is that they accept the message. In contrast to the beliefs of senders, beliefs about peer actions do not have an impact on receiver behavior.

REFERENCES

- Akerlof, G.A., 1970. The Market for 'Lemons': Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3), 488-500.
- Andreoni, J., Bernheim, B.D., 2009. Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5), 1607-1636.
- Ariely, D., Bracha, A., Meier, S., 2009. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review*, 99(1), 544-555.
- Battigalli, P., Charness, G., Dufwenberg, M., in press. Deception: The role of guilt. *Journal of Economic Behavior & Organization*.
- Bénabou, R. Tirole, J., 2006. Incentives and prosocial behavior. *American Economic Review*, 96(5): 1652-1678.
- Bicchieri, C., 2006. *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge University Press.
- Bicchieri, C., Xiao, E., 2009. Do the right thing: But only if others do so. *Journal of Behavioral Decision Making* 22(2), 191-208.
- Blume, A., DeJong, D.V., Kim, Y., Sprinkle, G.B., 2001. Evolution of communication with partial common interest. *Games and Economic Behavior* 37(1), 79-120.
- Brandts, J., Charness, G., 2003. Truth or consequences: An experiment. *Management Science* 49, 116-130.
- Cai, H., Wang, J.T., 2006. Overcommunication in strategic information transmission games. *Games and Economic Behavior* 56(1), 7-36.
- Cain, D.M., Loewenstein, G., Moore, D.A., 2005. The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *Journal of Legal Studies* 34(1), 1-25.

- Cain, D.M., Loewenstein, G., Moore, D.A., 2011. When sunlight fails to disinfect: Understanding the perverse effects of disclosing conflicts of interest. *Journal of Consumer Research* 37(5), 836-857.
- Charness, G., Dufwenberg, M., 2006. Promises and partnership. *Econometrica* 74(6), 1579-1601.
- Charness, G., Dufwenberg, M., 2010. Bare promises: An experiment. *Economics Letters* 107(2), 281-283.
- Childs, J., 2012.. Gender differences in lying. *Economics Letters* 114(2), 147-149.
- Crawford V.P., Sobel J., 1982. Strategic information transmission. *Econometrica* 50(6), 1431-1451.
- Dana, J., Weber, R.A., Kuang, J.X., 2007. Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1), 67-80.
- Dickhaut, J.W., McCabe, K.A., Mukherji, A., 1995. An experimental study of strategic information transmission. *Economic Theory* 6(3), 389-403.
- Dreber, A., Johannesson, M., 2008. Gender differences in deception. *Economics Letters* 99(1), 197-199.
- Ellingsen, T., Johannesson, M., 2004. Promises, threats, and fairness. *Economic Journal* 114(495), 397-420.
- Ellingsen, T., Johanneson, M., 2006. Pride and prejudice: The human side of incentive theory. *American Economic Review* 98(3), 990-1008.
- Ellingsen, T., Johannesson, M., Tjøtta, S., Torsvik, G., 2010. Testing guilt aversion. *Games and Economic Behavior* 68(1), 95-107.
- Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics* 114(3), 817-868.
- Fiedler, K., 1988. The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research* 50(2), 123-129.
- Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2), 171-178.
- Fosgaard, T.R., Hansen, L.G., Piovesan, M., in press. Separating will from grace: An Experiment on Conformity and Awareness in Cheating. *Journal of Economic Behavior & Organization*.
- Gibson, R., Tanner, C., Wagner, A.F., 2013. Preferences for truthfulness: Heterogeneity among and within subjects. *American Economic Review* 103(1), 532-548.
- Gneezy, U., 2005. Deception: The role of consequences. *American Economic Review* 95(1), 384-394.
- Gneezy, U., Erat, S., 2012. White lies. *Management Science* 58(4), 723-733.

- Greiner, B., 2004. An online recruitment system for economic experiments. In: Kremer, K. and Macho, V. (Eds.). *Forschung und wissenschaftliches Rechnen, GWDG Bericht 63*, Gesellschaft für wissenschaftliche Datenverarbeitung, Göttingen, 79-93.
- Hoffman, E., McCabe, K., Smith, V.L., 1996. Social distance and other-regarding behavior in dictator games. *American Economic Review*, 86(3), 653-660.
- Hurkens, S., Kartik, N., 2009. Would I lie to you? On social preferences and lying aversion. *Experimental Economics* 12(2), 180-192.
- Inderst, R., Rajko, A., Ockenfels, A., 2010. Transparency and disclosing conflicts of interest: An experimental investigation. Working paper, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie, Kiel.
- Kalai, E., Lehrer, E., 1995. Subjective games and equilibria. *Games and Economic Behavior* 8(1), 123-163.
- Kartik, N., Ottaviani, M., Squintani, F., 2007. Credulity, lies, and costly talk. *Journal of Economic Theory* 134(1), 93-116.
- Kartik, N., 2009. Strategic communication with lying costs. *Review of Economic Studies* 76(4), 1359-1395.
- Koch, C., Schmidt, C., 2010. Disclosing conflicts of interest - Do experience and reputation matter? *Accounting, Organizations and Society* 35(1), 95-107.
- Lacetera, N., Macis, M., 2008. Social image concerns and pro-social behavior. IZA discussion papers No. 3771.
- Li, M., Madarász, K., 2008. When mandatory disclosure hurts: Expert advice and conflicting interests. *Journal of Economic Theory* 139(1), 47-74.
- López-Pérez, R., Spiegelman, E., 2012. Do economists lie more? Working Papers in Economic Theory 2012/04, Universidad Autónoma de Madrid (Spain).
- López-Pérez, R., Spiegelman, E., 2013. Why do people tell the truth? Experimental evidence for pure lie aversion. *Experimental Economics* 16(3), 233-247.
- Lundquist, T., Ellingsen, T., Gribbe, E., Johannesson, M., 2009. The aversion to lying. *Journal of Economic Behavior and Organization* 70(1-2), 81-92.
- Ottaviani, M., Squintani, F., 2006. Naive audience and communication bias. *International Journal of Game Theory* 35(1), 129-150.
- Peeters, R., Vorsatz, M., Walzl, M., 2012. Beliefs and truth-telling: A laboratory experiment. University of Innsbruck, Working Papers in Economics and Statistics 2012-17.

- Popova, V., 2010. What renders financial advisors less treacherous? On commissions and reciprocity. *Jena Economic Research Papers* 2010-036.
- Reuben, E., Sapienza, P., Zingales, L., 2009. Is mistrust self-fulfilling? *Economics Letters* 104(2), 89-91.
- Rode, J., 2010. Truth and trust in communication: Experiments on the effect of a competitive context. *Games and Economic Behavior* 68(1), 325-338.
- Sánchez-Pagés, S., Vorsatz, M., 2007. An experimental study of truth-telling in a sender-receiver game. *Games and Economic Behavior* 61(1), 86-112.
- Sánchez-Pagés, S., Vorsatz, M., 2009. Enjoy the silence: an experiment on truth telling. *Experimental Economics* 12(2), 220-241.
- Stone, E.R., Yates, J.F., Parker, A.M., 1997. Effects of numerical and graphical displays on professed risk-taking behavior. *Journal of Experimental Psychology: Applied* 3(4), 243-256.
- Sutter, M., 2009. Deception through telling the truth?! Experimental evidence from individuals and teams. *Economic Journal* 119(534), 47-60.
- Vanberg, C., 2008. Why do people keep their promises? An experimental test of two explanations. *Econometrica* 76(6), 1467-1480.
- Weber, E.U., Siebenmorgen, N., Weber, M., 2005. Communicating asset risk: How name recognition and the format of historic volatility information affect risk perception and investment decisions. *Risk Analysis* 25(3), 597-609.

APPENDIX A

Sample screenshot decision making senders (translated from Spanish)

Please have a look at the following scenario:

There are three options that consist of a payoff for you and a payoff for player 2:

Option A : You receive **15** euros and player 2 receives **5** euros.
Option B : You receive **3** euros and player 2 receives **3** euros.
Option C : You receive **5** euros and player 2 receives **15** euros.

Please choose one of the following messages that should be sent to player 2 in case that this scenario will be implemented:

Message 1 : Option A will earn you more money than the other two options.
Message 2 : Option B will earn you more money than the other two options.
Message 3 : Option C will earn you more money than the other two options.

Choose one of the **messages (1, 2 or 3)** :

Figure A.1
Sample screenshot of the sender's decision in scenario 3.

APPENDIX B

Instructions for the experimental subjects (translated from Spanish)

Welcome to this experiment, we greatly appreciate your participation. Please switch off your cell phone. Communication with the other participants in the room is not allowed during the session. Read these instructions carefully, and raise your hand if you have any questions during the session. Your questions will be answered individually.

Your decisions in this experiment will allow you to earn a certain amount of money that we will pay you in cash at the end of the session.

You will be a player in a game that consists of two players. Your partner will be one of the other participants in this session, who the computer will randomly assign to you. Neither of you will know the identity of the other at any time. One of you will be assigned the role of "player 1" and the other the role of "player 2".

You will interact with this participant only once through the computers. After this interaction, the experiment will end, and you will be asked to fill in a short questionnaire.

Decision Making

During the experiment we will present to player 1 three scenarios with three options each. Each option consists of a payoff for player 1 and a payoff for player 2. The general structure of the options for each scenario that will be presented to player 1 is as follows:

Option A: Player 1 receives ... euros and player 2 receives ... euros.

Option B: Player 1 receives ... euros and player 2 receives ... euros.

Option C: Player 1 receives ... euros and player 2 receives ... euros.

We will present to player 1 the payoffs of both players for each option in each of the scenarios (the order of the options will be randomized). By contrast, player 2 will not be given this information. The task of player 1 is to choose one out of three messages that should be sent to player 2 afterward:

Message 1: Option A will earn you more money than the other two options.

Message 2: Option B will earn you more money than the other two options.

Message 3: Option C will earn you more money than the other two options.

Remember that there are three scenarios. This means that player 1 must decide, in each of the scenarios, which message she wants to be sent to player 2.

After player 1 has chosen a message for each scenario, the computer will randomly select one of the scenarios. This scenario will then be implemented, and the specific message that player 1 chose for this scenario will be sent to player 2. From this moment on, it depends on the decision of player 2 which of the three corresponding options will be implemented and, according to this decision, what amount of money both players will win.

Player 2 knows that there are three options available in the selected scenario, but she knows nothing about the profits associated with each option. The only information that player 2 receives is the message that player 1 chose for the implemented scenario.

After player 2 receives player 1's message, player 2 will make her decision, which is either to "accept" or "reject" the message. To "accept" the message means that player 2 accepts the information that is included in the message and that it is the option mentioned in the message that determines what both players will actually win. In contrast, to "reject" the message means that player 2 does not want the option mentioned in the message but another option to determine the profits for both of them. Therefore, if player 2 accepts the message, the option in the message will be implemented and determines the payoffs of the players. In the case that player 2 rejects the message, one of the two remaining options of the selected scenario will be implemented randomly by the computer to determine what each player wins.

Earnings

Before your payoffs are shown on the screen, you will answer some short questions. After that, player 1 will receive information about the acceptance or rejection of his message, the implemented option corresponding to the scenario that was selected by the computer and the payoffs for both players.

[Treatment T1:

However, player 2 will receive information only about her own payoff (corresponding to the implemented option). That means, player 2 will neither know her potential payoffs from the other options, nor the payoffs of player 1.]

[Treatment T2:

Player 2 will receive information about her own earnings corresponding to the implemented option. Furthermore, player 2 will receive information about all potential payoffs for both players in each of the options in the scenario that has been implemented.]

[Treatment T3:

In principle, Player 2 will only receive information about her own payoff corresponding to the implemented option. Furthermore, a possibility exists that the computer provides additional information to player 2 about all potential payoffs for both players in each of the options in the scenario that has been implemented. The probability of this happening is 50%.]

After that, we will pay you anonymously in cash the amount that corresponds to your final payoff in the game.

Do you have any questions about these instructions? If so, please raise your hand. If you do not have questions, remain silent until you get instructions from the experimenter.