WHEN DOES COMMUNICATION HARM?

VALERIA BURDEA*

July 8, 2020

Abstract

I investigate the effect of cheap talk in sender-receiver games where the sender wants the receiver to always take the same action, irrespective of the state of the world, while the receiver's preferences are state-dependent. These games model many economic situations such as investment, voting or purchase decisions. I consider environments where the interests of the players are either more likely aligned, or more likely conflicting. Using a simple theoretical analysis I show that communication can harm in settings where conflict of interests is less likely. This is due to increased skepticism resulting from receivers' aversion to being deceived ("sucker aversion"). However, when interests are more likely conflicting, communication can help due to senders' lying aversion. I run experiments to test these predictions and find that communication has a significantly positive effect on payoffs when interests are more likely conflicting. This cannot be explained by lying aversion only; social preferences seem to also matter. On the other hand, when interests are more likely aligned, opposite to previous findings, communication does not significantly increase receivers' skepticism and it does not affect payoffs.

Keywords: communication, sucker aversion, lying aversion, experiment

JEL Codes: D82, D83, D91

*University of Pittsburgh, Political Science Department and Pittsburgh Experimental Economics Laboratory (valeria.burdea@pitt.edu)

1 INTRODUCTION

Successful communication is one of the most important contributors to human betterment. Studies have shown that it can increase altruism (Mohlin and Johannesson, 2008), trust (Charness and Dufwenberg, 2006, 2011), coordination (Cooper et al., 1992) or cooperation (Bicchieri and Lev-On, 2007). But is communication a panacea or are there situations when one is better off turning a deaf ear?

This paper explores the role of communication in sender-receiver interactions with a sender that is better informed than the receiver about the payoff-relevant state of the world and also has state independent preferences. These games are of particular appeal as they model many economic situations such as investment, voting or purchase decisions. Consider the example of a seller who would like the buyer to buy his product irrespective of its quality or a politician who always wants the voter to prefer him rather than the opponent candidates. Such preferences have been given limited attention in the related literature due, perhaps, to the result that informative equilibria cannot arise in this setting (Crawford and Sobel, 1982). Nevertheless, from a behavioral perspective, such strategic settings can give rise to both positive and negative communication effects. In this paper I focus on lying aversion and trusting behavior as a source of positive communication effect. The negative effects arise from receivers' aversion to being profitably deceived or taken advantage of ("sucker aversion").

Sender's aversion to lie to the receiver is a behavioral characteristic that has been extensively studied in the experimental economics literature. Research has found that its roots could be an intrinsic cost of misrepresenting the truth (Abeler et al., 2019) or a cost stemming from guilty feelings for misleading the receiver (Charness and Dufwenberg, 2006). Irrespective of the source, evidence suggests that in strategic situations, senders behave as if they are averse to misrepresenting the nature of the world by transmitting more truthful information than standard equilibrium analysis would predict (Blume et al., 1998; Cai and Wang, 2006; Dickhaut et al., 2003; Forsythe et al., 1999). In this case, the receiver, whose interest is to find out the truth, could benefit from communication. This nevertheless depends on the receiver's likelihood to trust the truthful messages when there is no means to attest their veracity. In this respect, previous studies have shown that receivers are trusting enough in that they rely on the information provided by senders more than standard equilibrium analysis would predict (Cai and Wang, 2006; Kawagoe and Takizawa, 2009; Sánchez-Pagés and Vorsatz, 2009). However, whether the combination of lying aversion and trusting behavior are the channels that lead to efficiency-improving communication remains an open question.

When senders' preferences are state independent, their monetary incentives are usually perfectly transparent as they always prefer the receiver to take one action (e.g. buy the product, vote for them) rather than any other available one. This makes senders' incentives to deceive more salient for the receiver than if their preferred action would depend on the state of the world. Consequently, I hypothesize that an environment where senders have state-independent preferences is more likely to trigger a mistrusting behavior from the receiver in fear of the possibility of rewarding a deceiving sender. This tendency can make the receivers forgo profitable offers that would have otherwise been accepted in absence of communication. A similar behavior has been documented by Darke and Ritchie (2007) and Friestad and Wright (1994) pointing out that the simple knowledge that advertising tactics are designed to persuade buyers to buy a product can lead to increased consumer skepticism by making buyers more reluctant to buy the product.

Whether these hypothesized channels lead to welfare increasing or decreasing effects depends on the type of actions the receiver takes in absence of communication. There is little research to inform us on this topic. To the best of my knowledge, Ert et al. (2014) is the only paper that specifically tests for the effect of communication in a sender-receiver game by implementing a treatment without sender messages. The authors find that cheap talk makes the receiver more reluctant to choose the a-priori optimal action, but do not analyze whether communication leads to payoff losses in their experiment. Moreover, the question regarding why didn't a more trusting behavior take over is still un-addressed.

In this paper, I analyze a simple theoretical model where senders may be lying averse, i.e. they incur an additional, non-pecuniary cost from misrepresenting the truth, and receivers may be sucker averse, i.e. they incur an additional, non-pecuniary cost from trusting a potentially deceiving sender. Then, I show that outcomes can be driven by either sucker or lying aversion, as a function of the distribution of the states of the world. This means that communication can have both positive and negative effects, depending on the alignment of interests between sender and receiver. Specifically, when there is relatively less conflict of interests, communication can harm due to receiver's increased skepticism (a manifestation of sucker aversion). On the other hand, when there is relatively more conflict of interests, the lying aversion effect takes over, reducing buyer's skepticism and making both parties better off.

In this model, the likelihood of interests alignment between senders and receivers is a key modulator of trusting or skeptical receiver behavior. This is in line with recent evidence suggesting that individuals' propensity to lie depends on the probability of being considered a liar (Abeler et al., 2019; Gneezy et al., 2018). The latter is lower when the interests alignment between senders and receivers is higher as there are fewer incentives to lie, while the opposite is true when interests are more likely mis-aligned. Thus, when there is less conflict of interest, a sender who can benefit from lying is less likely to be considered a liar and therefore more likely to lie; when there is more conflict of interest, a sender who can benefit from lying is more likely to be considered a liar and therefore less likely to lie. In other words, the benefit of lying increases with a decrease in the conflict of interest between the sender and the receiver. Consequently, when interests are more likely to be aligned, communication leads receivers to be relatively more distrustful of the senders because the probability of lying is higher, while the opposite is true when interests are more likely opposed.

I implement a laboratory experiment to test these predictions using a one-shot seller-buyer game. In this game, a seller wants to sell a product to a buyer. The seller's product can be of high or low quality and the distribution of these states is common knowledge. The seller observes the randomly determined product quality and sends a message to the buyer. The buyer observes the message and decides whether to buy the product or not. Incentives are such that the seller would always like the buyer to buy the product. However, the buyer would like to buy only the high quality product. Therefore, the buyer's preferences depend on the state, while the seller's do not.

I use four treatments across which two dimensions are varied: whether the seller is allowed to send a free-form text message to the buyer and the frequency of high quality products which can take two values: $\frac{1}{3}$ and $\frac{2}{3}$. Note that this frequency essentially determines the conflict of interests between the players: $\frac{1}{3}$ defines an environment with high conflict of interests, while $\frac{2}{3}$ one with low conflict of interests. The results show that the effect of communication depends on the alignment of interest but not in the predicted manner. Specifically, I find that communication significantly increases trusting behavior in the $\frac{1}{3}$ - treatments while significantly increasing seller's average payoff, without decreasing the receiver's. Furthermore, in contrast with Ert et al. (2014), communication does not increase skepticism in the $\frac{2}{3}$ - treatments where it does not have a significant effect on players' payoffs.

This paper contributes to the literature on communication in several ways. First, it discusses when communication might be helpful as well as when it might be harmful, setting the stage for future research to further investigate the topic of when should one seek communication. Second, it conceptualizes a new potential mechanism through which communication can have negative effects: sucker aversion. Third, it introduces a simple experimental framework for the analysis of this effect of communication, that is flexible enough to accommodate a large variety of settings.

The rest of the paper is structured as follows: Section 2 gives a more detailed overview of the related literature, Section 3 introduces the theoretical model, analysis and derives testable hypotheses; Section 4 presents the experimental design and procedures; this is followed by the results analysis in Section 5 while Section 6 discusses the results and concludes.

2 RELATED LITERATURE

The paper is primarily related to experimental research on communication about private information. An important study in this field is that of Cai and Wang (2006) which implements treatments to test the predictions of the Crawford and Sobel (1982) model of strategic information transmission (cheap talk). In their game, players are paired and assigned to the role of sender or that of receiver. The sender has private information about the state of the world, s. The value of s is a number uniformly distributed over the state space $S = \{1, 3, 5, 7, 9\}$. They use strict-form messages which can take any value from $M = \{1, 3, 5, 7, 9\}$. After observing the message, the receiver takes an action from $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. The receiver wants to take an action as close as possible to the state while the sender's preferred action is $s + d$, where d represents the preference difference (size of the conflict of interests). They implement treatments with different values for d and find support for the theoretical predictions according to which as d increases, less information is transmitted by senders and used by receivers. As opposed to equilibrium predictions, they observe that senders report the state truthfully more than they should (over-communication). This result is in line with Dickhaut et al. (1995) who investigate a similar topic in a similar environment. The current study differs from these papers in several ways. First, in my game the sender has state-independent preferences and therefore the parameter d is not relevant. Second, I replace the concept of deterministic preference difference with that of the likelihood that preferences differ by varying the receiver's optimal action (defined as such from a standard theoretical point of view). Third, I implement treatments also without communication so that I can compare behavior with an empirical benchmark as well, not only with a theoretical one.

Blume et al. (1998, 2001) also investigate the strategic transmission of information in an experimental setting and their results are largely in line with the over-communication phenomenon documented by Cai and Wang (2006), However, in their studies the focus is on how the meanings of messages evolve giving rise to a different setting than the one we are investigating in this chapter. Nevertheless, these results support the hypothesis that lying aversion is a robust phenomenon in strategic frameworks.

Another relevant paper is Forsythe et al. (1999) who investigate a buyer-seller setting characterized by inefficient trading due to adverse selection (Akerlof, 1970). They test whether two different communication mechanisms lead to efficiency gains (more trade). In the first mechanism, "cheap talk", sellers' message can be any subset of product qualities (not necessarily the true one). In the other mechanism, "antifraud", sellers' report must include the true quality. They find that both mechanisms increase trade efficiency but in the cheap talk environment this gain comes at the buyer's expense. Our studies are similar in that the seller's preferences are state-independent, while the buyer's are not. Nevertheless, we differ in several dimensions. In their experiment, buyers and sellers with different valuations for an asset have to trade in a market by bidding prices. Therefore, unlike in my setting, sellers are active players even when communication is not possible. Also, they focus only on a framework where the different asset qualities are equally likely while I explore scenarios in which the distribution is not uniform.

The closest paper to my study comes from the negotiation literature where Ert et al. (2014) investigate behavior in a cheap-talk game with a seller who has state-independent preferences. In their game, the seller randomly draws two cards from a deck containing 100 cards with consecutive values from 1 to 100. The buyer is then informed about the value of the lower card and decides whether to buy the two cards from the seller or not at a fixed price. The seller's payoff is the fixed price while the buyer's payoff is represented by the sum of the two cards minus the fixed price. In their experiments, the authors fix the value of the lower card to 40 which makes it optimal for the buyer to buy the cards[1]. They compare treatments with and without communication to investigate whether communication makes buyers more skeptical to buy and find a significant skepticism effect. In the present study, I simplify the setting by implementing a cheap-talk game with only 2 possible states. This should reduce the noise that free-form communication might introduce. Second, I am interested in studying the evolution of skepticism when the buyer's optimal action is different than the seller's preferred action. In the terms of Ert et al. (2014) this is equivalent to the buyer's ex ante optimal action being not to buy.

Serra-Garcia et al. (2011) investigate the effect of vague and precise communication in a sequential 2-player public good game with asymmetric information about the value of the public good. This can be high, intermediate or low. Like in the present paper, they hypothesize that communication can decrease efficiency, though their mechanism is based on the leader's lying aversion rather than on the follower's skepticism. They do not find support for this hypothesis as leaders frequently lie with precise communication, while when messages can be vague, followers do not correctly interpret them.

Other similar game environments are found in the literature on deception detection. Belot and Van de Ven (2017) also use a buyer-seller setting with asymmetric information. The main difference with my study is the communication protocol. Specifically, Belot and Van de Ven (2017) are interested in whether buyers can spot deception in face-to-face interactions. I, on the other hand, am interested in whether anonymous communication can improve outcomes compared to a benchmark in which communication is not allowed. Nonetheless, the results of this study are relevant as the ability to spot deception can be a mechanism that leads to gains from communication. The authors find that buyers are better than chance in predicting which sellers are lying when they interact through face-to-face communication. This result supports the finding of Belot et al. (2012) who study the ability of people to detect deceit during a television game show where players are competing in a high-stakes, simultaneous prisoner's dilemma game. Other studies using face-to-

[1] They do so by using two rounds: one in which the lower card's value is 40, and one in which it is randomly drawn. The subjects are not informed about the value of the lower card but only that in one of the two rounds this value was pre-determined.

face interaction also find that subjects are better than chance when asked to predict the behavior (Brosig, 2002; Dawes et al., 1977) or private information of others (Ockenfels and Selten, 2000). Deceit-detection in written communication, such is the case in my study, is though less explored in the literature. Chen and Houser (2017) investigate whether observers of a three-person trust game can predict the trustworthiness of free-form written messages. They find little support for subjects' ability to predict better than chance.

3 THEORETICAL ANALYSIS

In this section I introduce the game and the theoretical predictions according to standard theory. Then, I explore how these predictions might change if we consider lying averse sellers who incur a psychological cost from lying and buyers who are averse to being profitably deceived. Such buyers incur a psychological cost from taking an action which rewards a deceiving seller. We will call this cost a sucker cost. Furthermore, I investigate the role of these behavioral biases when incentives are more likely to be aligned compared to when it is more probable that they are unaligned.

3.1 *The game*

In the benchmark case of our experiment, the following game is played between two players: a seller and a buyer. The timeline of the game is the following:

Stage 1: Nature's move. *Nature* randomly determines the state of the world (s), represented by a ball that can take one of two possible colors: red (R) or blue (B). Hence, the state space is $S = \{R, B\}$. The state s is private information of the seller. The probability of $s = R$ is p and this is common knowledge; $p \in [0, 1]$.

Stage 2: Seller's move. After observing the state, the seller can send a free-form message (m) to the buyer. For the purpose of this analysis we will restrict the message space to two messages: reporting that the ball is red and reporting that the ball is blue. The message space under consideration is therefore $M = \{R, B\}$.

Stage 3: Buyer's move. After receiving the message from the seller, the buyer decides whether to buy the red product or the blue product. This is equivalent to taking an action (a) from the following set: $A = \{R, B\}$.

The payoff of the seller depends only on the buyer's action while the buyer's payoff depends both on his action and on the true state of the world. Table 1 summarizes players' material payoffs.

Table 1: Payoff Matrix (seller's payoff listed first in each cell)

	$a = R$	$a = B$
$s = R$	(1, 1)	(0, 0)
$s = B$	(1, 0)	(0, 1)

The seller, therefore, would always like the buyer to buy the red product (state-independent preferences) while the buyer would like to buy the red product only when the state is red and to buy the blue product when the state is blue. We assume that messages have literal meanings. In a series of experiments Blume et al. (2001) show that indeed, players tend to use messages with their natural language interpretations.

3.2 *Equilibrium analysis*

The solution concept is Perfect Bayesian Equilibrium (PBE); the analysis is restricted to pure strategies. A strategy for the seller is a vector $m = (m(R), m(B))$ where $m(R)$ represents the seller's message decision when the state is red, and $m(B)$ his message decision when the state is blue. A strategy for the buyer is a vector $a = (a(R), a(B))$ where $a(R)$ represents the buyer's action when the message is red, and $a(B)$ his action when the message is blue. Moreover, let $\mu(m) \in [0, 1]$ be the posterior probability that the buyer assigns to $s = R$ upon observing m. All players are risk neutral, material payoff maximizers. A PBE of the game requires (1) sequential rationality of each players' strategy - at any information set, a player uses a best response strategy given their beliefs and holding the other player's strategy constant; (2) consistency of beliefs - each player's beliefs follows Bayes' rule (wherever appropriate) and is consistent with the strategy profile

It is easy to see that the buyer's ex ante optimal action when $p > \frac{1}{2}$ is $a = R$, whereas when $p < \frac{1}{2}$ it is $a = B$. When $p = \frac{1}{2}$, the buyer is indifferent between the two actions and we will assume that they choose $a = R$ in this case. When introducing messages, it is useful to distinguish between three types of equilibria: uninformative, informative and persuasive ones. In an uninformative equilibrium, the buyer's beliefs and actions are the same as when communication would not be possible. In an informative equilibrium the buyer's belief about the probability that $s = R$ conditional on the message is higher than the prior but not high enough to make the buyer choose an action different than his ex ante optimal one. Lastly, in a persuasive equilibrium, the buyer's optimal action is different than his ex ante one in a way that the seller's payoff is higher than in any uninformative equilibrium. Note also that uninformative or informative equilibria can also be characterized as unpersuasive equilibria.

In a recent paper, Lipnowski and Ravid (2017) prove that in a cheap-talk game with two buyer actions, two states and sender state independent preferences where the buyer is not indifferent between the two actions at the prior, informative equilibria exist but persuasive ones are impossible. The following two lemmas follow trivially.

Lemma 1. *In every PBE of the game where $p > \frac{1}{2}$, seller's message strategy (m) is not persuasive and $a(R) = a(B) = R$, on the equilibrium path.*

Lemma 2. *In every PBE of the game where $p < \frac{1}{2}$, seller's message strategy (m) is not persuasive and $a(R) = a(B) = B$.*

Note that when $p > \frac{1}{2}$, it is relatively more likely that the interests of the players are aligned and the buyer's optimal action coincides with the seller's preferred one. The opposite is true when $p < \frac{1}{2}$. Any seller strategy can be supported in equilibrium in both cases. However, if messages have a natural language interpretation (Blume et al., 2001), it is reasonable to assume that sellers would be more likely to signal their preferred action, i.e. red, when the state coincides with this action, rather than signaling that the state is blue.

3.3 *Introducing lying and sucker costs*

3.3.1 *Lying cost*

We begin with a simple analysis of players' behavior when sellers are potentially lying averse. We will call this game "the L-game". For simplicity, we assume that there are two types of sellers: liars (L) and truth-tellers (T). Liars are material payoff maximizers. Truth-tellers are assumed to be lying averse in that they incur a psychological cost from sending a message $m \neq s$. This cost is high enough such that they always tell the truth.

We implement this in our game by letting nature decide not only the state but also the seller's type. Let $\lambda \in [0, 1)$ be the proportion of liars in the population of sellers. We assume that λ is common knowledge. Therefore, the buyer does not know whether he is interacting with a T or L seller but he knows the distribution of these two types in the population of sellers. The case where $\lambda = 1$ is equivalent to one in which sellers are selfish rational utility maximizers and the analysis is the same as in Section 3.2. When $\lambda = 0$, all sellers are truth-tellers and the buyer's optimal strategy is to follow the message.

There are four possible pure strategies for the L seller:

(a) $(m_L(R), m_L(B)) = (R, R)$

(b) $(m_L(R), m_L(B)) = (B, B)$

(c) $(m_L(R), m_L(B)) = (R, B)$

(d) $(m_L(R), m_L(B)) = (B, R)$

The T seller always chooses $(m_T(R), m_T(B)) = (R, B)$. Next, we check which of the potential L seller strategies can be supported in a PBE of the game and what is the corresponding receiver best reply.

(a)

Lemma 3. *The following strategies:* $(m_L(R), m_L(B)) = (R, R)$, $(a(R), a(B)) = (R, B)$, *are part of a PBE of the L-game where* $p \geq \frac{1}{2}$. *When* $p < \frac{1}{2}$, *they are a PBE if and only if* $\lambda \leq \frac{p}{1-p}$. *If* $\lambda > \frac{p}{1-p}$, $(m_L(R), m_L(B)) = (R, R)$, $(a(R), a(B)) = (B, B)$, *are part of a PBE of the L-game.*

Proof. See Appendix.

□

(b)

Lemma 4. *The following strategies:* $(m_L(R), m_L(B)) = (B, B)$, $(a(R), a(B)) = (R, R)$ *are part of a PBE of the L-game if and only if* $p \geq \frac{1}{2}$ *and* $\lambda > \frac{1-p}{p}$. *If* $p < \frac{1}{2}$, $(m_L(R), m_L(B)) = (B, B)$ *cannot be supported in equilibrium.*

Proof. See Appendix.

□

(c)

Lemma 5. *The following seller strategy:* $(m_L(R), m_L(B)) = (R, B)$, *cannot be supported in a PBE of the L-game.*

Proof. See Appendix.

□

(d)

Lemma 6. *The following strategies:* $(m_L(R), m_L(B)) = (B, R)$, $a(R) = a(B)$ *are part of a PBE of the L-game where* $p \geq \frac{1}{2}$ *if and only if* $p > \lambda > 1 - p$ *or* $\lambda > p$. *The strategies form a PBE when* $p < \frac{1}{2}$ *if and only if* $\lambda > 1 - p$ *or* $\lambda < p$.

Proof. See Appendix.

□

The L-game, therefore has both persuasive and unpersuasive equilibria. To study the effect of communication on players' payoffs I will focus on persuasive equilibria. The L-game has only one persuasive equilibrium described in Lemma 3.

The L seller's expected payoff in the persuasive equilibrium of the L-game is equal to 1 when $p \geq \frac{1}{2}$. This is the same as the seller's payoff in a no-communication protocol (equivalent to the case described by Lemma 1). The buyer's payoff in the persuasive equilibrium is equal to $p + (1 - p)(1 - \lambda)$. This can be higher than p (the buyer's no-communication expected payoff) depending on the proportion of T sellers $(1 - \lambda)$. The T seller's expected payoff in this case is equal to p which is lower than the L seller's expected payoff.

Proposition 1. *In the L-game with $p \geq \frac{1}{2}$, if $\lambda < 1$, the buyer's expected payoff in a persuasive equilibrium is higher than his expected payoff in a no-communication protocol. The seller is at most as well off with communication than without.*

When $p < \frac{1}{2}$, the L-seller's expected payoff in the persuasive equilibrium of the L-game is again equal to 1. This is higher than the L-seller's payoff in a no-communication protocol (described by Lemma 2) which is equal to 0. The buyer's expected payoff (equal to $p + (1 - p)(1 - \lambda)$) is higher than their no-communication expected payoff (p) as long as $\lambda < 1$. This is automatically satisfied by the condition for the equilibrium ($\lambda \leq \frac{p}{1-p}$) when $p < \frac{1}{2}$. The T seller's expected payoff is $\frac{1}{3}$ which is also greater than if communication were not possible. The buyer's expected payoff is equal to $(1 - p)(1 - \lambda) + p$ which is greater than $1 - p$ given that $\lambda \leq \frac{p}{1-p}$. Proposition 2 follows from this analysis.

Proposition 2. *In the L-game with $p < \frac{1}{2}$, if $\lambda \leq \frac{p}{1-p}$, all players' expected payoffs in a persuasive equilibrium are higher than their expected payoffs in a no-communication protocol.*

3.3.2 *Sucker cost*

We now introduce non-standard preferences for the buyers. Specifically, we assume that the buyer can incur a sucker cost whenever he chooses $a(R) = R$ after a deceptive message ($m = R$ when $s = B$). When the buyer takes an action, he does not know whether the message is deceptive or not but has an expectation about this. The higher this expectation, the more likely he is to incur the sucker cost.

Again, we restrict our analysis to only two types of buyers: those with a high enough cost which we will call "skeptics" (K) and those whose cost is zero which we will call "non-skeptics" (N). Specifically, K buyers always choose $(a_K(R), a_K(B)) =$

(B, B), whereas N buyers best reply to each message. Let $p_k \in [0, 1]$ be the proportion of skeptics in the population of buyers. This, in essence, represents the probability that the buyer is a skeptic. We will call this game "the K-game". Next, we analyze the equilibria of this game given each parametrization.

Lemma 7. *In every PBE of the K-game where $p \geq \frac{1}{2}$ and $p_k < 1$, on the equilibrium path, $(a_N(R), a_N(B)) = (R, R)$ while $(a_K(R), a_K(B)) = (B, B)$.*

Proof. See Appendix. □

When $p < \frac{1}{2}$, nothing changes in the equilibrium predictions compared to the original game. This is because the equilibrium when $p < \frac{1}{2}$ is that $a(R) = a(B) = B$ and so the N buyer and the K one behave the same.

Based on the previous results we formulate the following prediction:

Proposition 3. *In the K-game where $p \geq \frac{1}{2}$, communication leads to material losses for the seller and the K buyer if $p_k > 0$. Communication has no effect when $p < \frac{1}{2}$.*

Proof. The proof for why communication has no effect in the $p < \frac{1}{2}$ case of the S-game is trivial and it follows from Lemma 2, and the fact that behavior doesn't change when we account for p_k in this setting. To see why communication leads to losses when $p \geq \frac{1}{2}$, consider the players' expected payoffs in the S-game:

$$\pi_{K_{buyer}} = 1 - p$$

$$\pi_{seller} = 1 \cdot (1 - p_k) + 0 \cdot p_k = 1 - p_k$$

Recall that the buyer's expected payoff in a no-communication protocol is p, while the seller's is 1. Therefore, if $p > \frac{1}{2}$ and $p_k > 0$, both the K buyer's and the seller's expected payoffs in the S-game are smaller than their expected payoffs in a no-communication protocol.

□

Proposition 3 implies that as long as skeptics exist, both the average buyer and the seller should lose from communication in the $p \geq \frac{1}{2}$ case, but not in the $p < \frac{1}{2}$ one.

3.3.3 *Lying cost + sucker cost*

Finally, we look at the implications of the situation where both the seller and the buyer have non-standard preferences. In particular, the seller has a lying cost as in Section 3.3.1 while the buyer has a sucker cost as in Section 3.3.2. We will call this the "LK-game".

Notice that nothing changes in the equilibrium analysis from Section 3.3.1 when considering the possibility of skeptics, as long as $p_k < 1$. This is because, in any equilibrium in which the non-skeptic buyer would optimally choose $a = R$, the non-lying averse seller has no incentive to deviate from their strategy as the skeptic buyer always chooses $a = B$. What is the payoff effect for the average buyer and seller?

Proposition 4. *In the LK-game where $p \geq \frac{1}{2}$, if and only if $p_k > \frac{(1-\lambda)(1-p)}{p+p\lambda-\lambda}$ and $p_k > 0$, communication leads to lower average payoffs for both players. If $0 < p_k < \frac{(1-\lambda)(1-p)}{p+p\lambda-\lambda}$, communication helps the average buyer while harming sellers.*

Proof. The average buyer's expected payoff in the LK-game when $p \geq \frac{1}{2}$ is $p_k(1-p) + (1-p_k)(p + (1-p)(1-\lambda))$. This is lower than p, their expected payoff without communication, as long as $p_k > \frac{(1-\lambda)(1-p)}{p+p\lambda-\lambda}$. Note that this is true irrespective of λ when $p > \frac{1}{2}$ since in this case, $\lambda < \frac{p}{1-p}$.

The average seller's expected payoff in the communication game is $p(1-p_k) + (1-p)\lambda(1-p_k)$. Their expected payoff in the no-communication game is 1. Therefore, as long as $p_k > 0$ or $\lambda > 0$, the average seller's expected payoff in the LK-game is lower than in the no-communication game. □

The intuition behind this result is that the average buyer's gains from interacting with lying averse sellers plus those from being skeptical about liars compensate his losses from being skeptical towards truth-tellers as long as there is not too much skepticism in the population. If skepticism is low enough, the gains from interacting with lying averse sellers take over. For the seller, however, both lying aversion and skepticism lead to negative payoff effects irrespective of the distribution of skeptics in the population. What happens when $p < \frac{1}{2}$?

Proposition 5. *In the LK-game with $p < \frac{1}{2}$, if and only if $\lambda < \frac{1}{2}$ and $\lambda < \frac{p}{1-p}$, both players can gain from communication.*

Proof. If $\lambda > \frac{1}{2}$, seller's messages are uninformative and the only equilibrium is the one where every buyer chooses $a = B$ irrespective of message. Hence, in this case communication makes no difference. If $\lambda < \frac{1}{2}$, the average buyer's expected payoff in the LS-game when is $p_k(1-p) + (1-p_k)(p + (1-p)(1-\lambda))$. In order for communication to help, this payoff has to be higher than the buyer's expected payoff in the no-communication game, which is $1-p$. This reduces to the condition: $p + k(\lambda - p\lambda - p) > \lambda - p\lambda - p$. Note that when $\lambda - p\lambda - p > 0$ which means that $\lambda > \frac{p}{1-p}$, the condition reduces to $p_k > 1$, which is not possible. In other words, if everyone is a skeptic, communication plays no role. However, when $\lambda < \frac{p}{1-p}$, the for communication to help the average buyer reduces to $p_k < 1$.

The average seller's expected payoff in the LS-game where $\lambda < \frac{1}{2}$ and $p < \frac{1}{2}$ is $p(1-p_k) + (1-p)\lambda(1-p_k)$. Their expected payoff in the no-communication game is 0. Therefore, as long as $p_k < 1$ and $\lambda > 0$, the average seller's expected payoff in the LK-game is higher than in the no-communication game.

□

4 EXPERIMENTAL DESIGN

4.1 *Procedures*

The experiment was conducted at the Centre for Decision Research and Experimental Economics (CeDEx) laboratory located at the University of Nottingham. The decision-making environment was computerized and the software was programmed in z-Tree (Fischbacher, 2007). There were 464 participants (out of which 270 were females) recruited from a university-wide pool of undergraduate and graduate students using ORSEE (Greiner, 2015). Participants were randomly assigned to one of the 16 sessions. Each session involved between 26 and 30 participants and lasted on average 20 minutes. All subjects received a £3 participation fee plus their earnings from playing the game. The game earnings amounted to £2 on average. In the experiment, payoffs were described in points (30 or 0) which were then exchanged to pounds according to the following exchange rate: 1 point=£0.1.

I used a 2x2 design varying the following two dimensions: the possibility of sellers to send messages and p. Essentially, what p represents is the likelihood that the interests of the seller and the buyer are aligned. These values were chosen to reflect a case where interests are more likely to be aligned ($p = \frac{2}{3}$) and one where they are more likely to be conflicting ($p = \frac{1}{3}$). Table 2 presents a summary of the 4 treatments.

Table 2: Treatments summary

Name	Messages allowed?	Value of p
CM_23	YES	$\frac{2}{3}$
No_CM_23	NO	$\frac{2}{3}$
CM_13	YES	$\frac{1}{3}$
No_CM_13	NO	$\frac{1}{3}$

There were 4 sessions for each treatment. Upon arriving at the lab, participants were randomly assigned a number that would determine their seat in the lab. At their computer terminals, subjects would find a sheet of instructions[2] which were read out loud by the experimenter once everyone was seated. After the instructions were read, the experiment began with a set of hypothetical control questions regarding subjects' understanding of the payoff matrix. The experiment then proceeded with the game relevant for their respective treatment. Participants played the game only once. Finally, they were asked to fill in a short questionnaire after which they were paid in private and in cash.

Our communication protocol was computerized, free-form and one-way in that only the seller could send a message. The seller's screen is depicted in figure 1. I used free form messages for comparability with Ert et al. (2014) but also to create an environment more likely to promote skepticism.

[2] See Appendix for a full set of instructions.

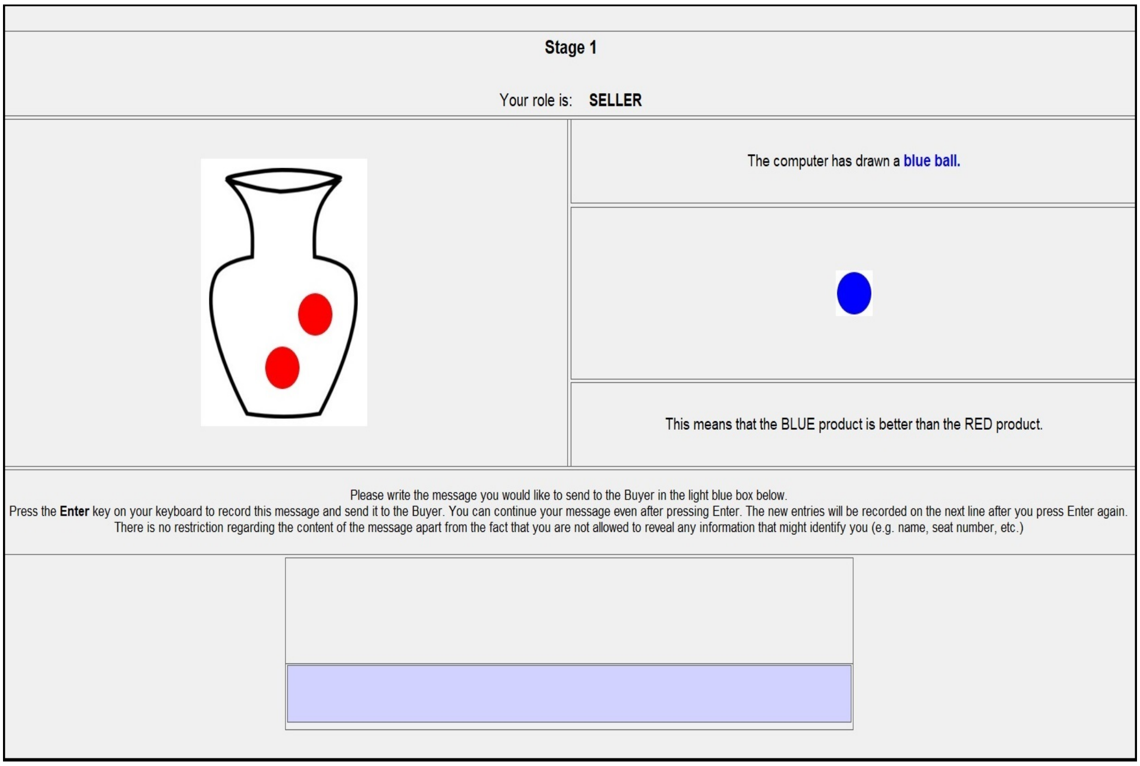


Figure 1: Communication interface

The state of the world was randomly determined for each subject at the start of the experiment. The exact distribution of states in each treatment is presented in Table 3.

Table 3: No. of observed red and blue states across treatments

Treatment	$s = red$	$s = blue$
CM_23	37	21
No_CM_23	42	16
CM_13	22	38
No_CM_13	19	37

4.2 *Hypotheses*

In this section I formulate hypotheses regarding behavior that draw from the previous theoretical analysis. I note to which type of assumption regarding players' preferences each hypothesis refers to.

Standard preferences:

The following two hypotheses are based on the game in which players have selfish, rational preferences.

Hypothesis 1. *Buyers will always choose $a = R$ when $p = \frac{2}{3}$ and $a = B$ when $p = \frac{1}{3}$, irrespective of whether communication is present.*

Hypothesis 2. *Communication has no effect on material payoffs irrespective of p.*

The second hypothesis is an implication of the fact that buyers' actions do not depend on messages, and so their presence does not affect final outcomes.

Non-standard preferences:

The following hypotheses refer to the L-game only, when there are lying-averse sellers but no sucker-averse buyers.

Hypothesis 3. *When $p = \frac{2}{3}$, and a positive fraction of sellers report blue states truthfully, communication improves buyers' average material payoff and decreases the sellers' one. Communication has no effect on outcomes in this case when no truth-telling is observed.*

Hypothesis 4. *When $p = \frac{1}{3}$, and more than half of sellers report blue states truthfully, communication improves the average material payoffs of both players. Communication has no effect on outcomes in this case when less than half of sellers report blue states truthfully.*

Next, we account for the existence of skeptical buyers.

Hypothesis 5. *When $p = \frac{2}{3}$, the share of buyers buying the blue product after communication is higher than the share of buyers choosing to buy the blue product when messages are not allowed. When $p = \frac{1}{3}$ and $l < \frac{1}{2}$, the share of buyers buying the red product after communication is higher than the share of buyers buying the red product when messages are not allowed.*

Hypothesis 6. *When $p = \frac{2}{3}$, communication reduces the average payoffs of both players. When $p = \frac{1}{3}$, communication increases the average payoffs of both players.*

Finally, given the hypothesized (negative) effect of communication in the $p = \frac{2}{3}$ case and its counterpart in the $p = \frac{1}{3}$ case we can formulate a general hypothesis regarding the average effect of communication.

Hypothesis 7. *Communication has an absolute effect on outcomes.*

5 RESULTS

In this section I analyze the behavior employed by buyers and seller across treatments and compare it with our hypotheses. I then investigate the effect of this behavior on payoffs. Throughout the analysis I will compare the observed outcomes both between treatments as well as with the standard equilibrium benchmark. Since players played the game only once and there was no interaction across the buyer-seller pairs, each subject's decision represents an independent observation.

5.1 *Seller behavior*

This section looks at sellers' message strategy in the CM_ treatments as only there the sellers are active players. Since messages are free-form, I classify them into well-defined categories. The free-form message space can be partitioned in three types to accommodate the meanings of all observed messages. Table 4 presents the identified types, an example for each category and their distribution within each treatment.

Table 4: Message types and corresponding frequencies

Type	Example	Frequency	
		CM_23	CM_13
RED	"Hey, I drew the red ball - the red product is better than blue!"	85%	70%
BLUE	"blue ball was drawn so pick the blue product"	9%	12%
??	"This product is good quality compared to the other in the urn."	6%	18%

First, we note the use of a third type of message, the inconclusive one, denoted by '??' in the above tables. This type of message does not make a claim about the color of the state. Its use could be explained if inconclusive messages are an indirect form of lying (evasive lying). Using such a message when the state is blue could give rise to lower costs than direct lying (Khalmetski et al., 2017). If this is the case, we should observe their use only when the state is blue. Figure 2 presents how messages are distributed across red and blue states for each communication treatment.

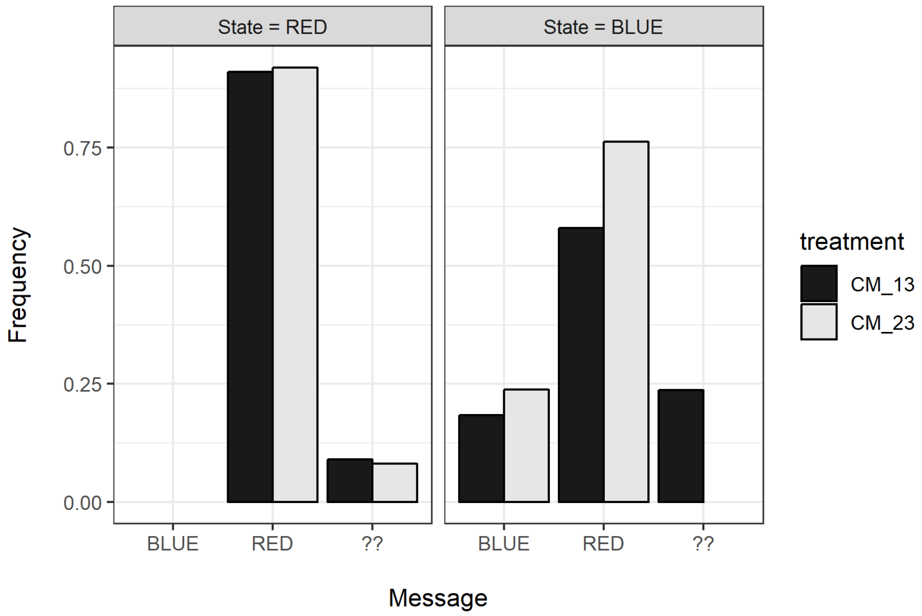


Figure 2: Message distribution across states and treatments

From figure 2, we observe that no seller reports red states as blue. This is an indication that messages are used with their literal meanings. In terms of truthfulness of messages, we note that both in the CM_23 treatment and in the CM_13 one, a significantly positive proportion of sellers report blue states truthfully: 23.8% in CM_23 and 18.4% in CM_13. These frequencies are both significantly different than 0 ($p - value = 0.010$ for CM_23 and $p - value = 0.003$ for CM_13, Barnard's one-sided test[3]) but not significantly different compared to each other. This represents empirical support for the existence of lying averse sellers. Their lying cos, though, does not seem to be affected by the probability of being considered a liar (which varies across treatments).

Result 1. *A significant but low proportion of sellers report blue states truthfully. This proportion is not different across the CM_ treatments.*

The proportion of lying averse sellers in this experiment is rather small compared to typical results regarding truth-telling propensity in individual-decision making environments. Abeler et al. (2019) estimate an average amount of truth-telling at around 78%. An explanation for why the observed lying frequency in the current study is so small could be the fact that strategic environments have been found to crowd out lying aversion (Cabrales et al., 2020; Minozzi and Woon, 2013).

[3]I also perform the Fisher test and Chi-squared test by Monte Carlo simulation for all results in this section and the findings are qualitatively the same.

The proportion of truth-tellers in our sample is quite similar to the 22% one observed by Sánchez-Pagés and Vorsatz (2007) in a sender-receiver game. It is also too low to observe positive effects from communication as per Hypothesis 5.

How is the probability of sending a red message affected by the treatment? Table 5 presents the marginal effects resulting from probit regressions of the message choice on a treatment variable and control variables (gender and age). I run separate regressions depending on the underlying state to obtain a more accurate picture of the meaning of different messages and how this might change across treatments.

Table 5: Probit marginal effects of message decision

	Dependent variable: m=RED		*Dependent variable: m=??*	
	state=RED	state=BLUE	state=RED	state=BLUE
CM_23	0.009	0.164	−0.009	−0.191***
	(0.076)	(0.130)	(0.076)	(0.073)
Female	−0.001	−0.284**	0.001	0.011
	(0.074)	(0.124)	0.074	(4.938)
Age	−0.055	−0.002	0.005	0.000
	(0.012)	(0.035)	(0.012)	(0.072)
Observations	59	59	59	59
Note:				*p<0.1; **p<0.05; ***p<0.01

The results show that there is no significant treatment effect on the probability of sending the red message irrespective of the value of the state. We observe a significantly negative gender effect on the probability of sending the red message when the state is blue. In other words, females are 28% less likely to lie by sending a red message. This is a typical finding in the deception literature (Capraro, 2018).

Furthermore, the regressions show that the likelihood of sending an inconclusive message is significantly different across treatments when the state is blue (in which case there is an incentive to lie). Specifically, being in the CM_23 treatment significantly decreases the probability of sending an inconclusive message with 19.10%. This is an indication that the meaning of inconclusive message differs across treatments. Nevertheless, this result should be taken with a grain of salt as the frequency of these messages is overall low.

5.2 *Buyer behavior*

Figure 3 depicts how often buyers choose their ex ante optimal strategy across treatments. Recall that when $p = \frac{2}{3}$, the ex ante optimal action for the buyer is to buy the red product, whereas when $p = \frac{1}{3}$, it is to buy the blue product.

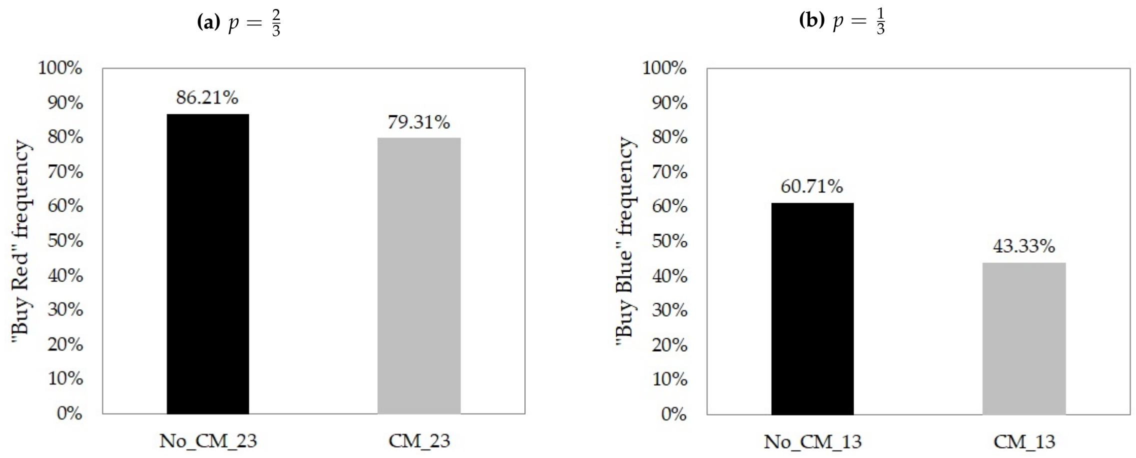


Figure 3: Frequency with which buyers choose the ex ante optimal action

First, we notice that communication makes buyers choose the ex ante optimal action less often. This difference is (weakly) significant in the $p = \frac{1}{3}$ case ($p-value = 0.064$, Barnard's two-sided test). To better understand this finding we look at how buyers' choice of action depends on the message received. Table 6 reports the distribution of buyer's actions conditional on the message received.

Table 6: Distribution of buyer actions conditional on the message received

(a) CM_23

	$a = red$	$a = blue$
$m = red$	87.75%	12.24%
$m = blue$	20.00%	80.00%

(b) CM_13

	$a = red$	$a = blue$
$m = red$	73.81%	26.19%
$m = blue$	14.29%	85.71%

We observe a significantly high proportion of buyers who choose $a = red$ when $m = red$ in both treatments. If we focus on the cases where $m = red$, the effect of communication is even stronger when $p = \frac{1}{3}$, the frequency of $a = red$ being 73.8%, which is significantly higher than 39.3% ($p - value = 0.001$, Barnard's two-sided test).

Result 2. *Buyers choose the ex ante optimal action less often in the CM_ treatments than in the No_CM_ ones. This difference is significant only in the $p = \frac{1}{3}$ case.*

The corollary of Result 2 is that communication is persuasive in the CM_13 treatment as it makes buyers take the action preferred by the seller more often. By analyzing buyers' behavior conditional on the messages received we can also find out whether communication increases buyers' skepticism in the $p = \frac{2}{3}$ case as hypothesized. We observe that the percentage of buyers choosing to buy the blue product after a red message in the CM_23 treatment is equal to 12.24%. This is significantly different from 0 ($p - value < 0.010$, one-sided binomial test). Is this evidence for sucker aversion? Recall that the validity of this claim depends on skepticism being observed only in the presence of messages. Therefore, our model implies that in the No_CM_23 treatment, this frequency would be equal to 0. Nonetheless, we find this is not the case as the frequency with which buyers buy the blue product in the no communication treatment is 13.8%. Therefore, the reason why buyers buy the blue product after a red message in the _23 treatments is not because of sucker aversion.

Result 3. *Communication does not increase skepticism when $p = \frac{2}{3}$ and therefore, the observed deviation from the ex ante optimal strategy is not due to sucker aversion.*

Result 3 goes against Hypothesis 5. Why do we only observe a communication effect in the $p = \frac{1}{3}$ case? Focusing on the No_CM_ treatments, an interesting finding arises: buyers choose their ex-ante optimal action significantly less often than equilibrium predicts ($p - value = 0.004$ in No_CM_23 and $p - value < 0.001$ in No_CM_13, Barnard's two-sided test).

Result 4. *Buyers' behavior is significantly different than standard equilibrium predictions in the No_CM_ treatments.*

How is the probability of choosing the ex-ante optimal action affected by communication and the conflict of interests? Table 7 presents the results of a linear regression of the buyer's decision being ex-ante optimal (i.e. red, when p=$\frac{2}{3}$ and blue when p=$\frac{1}{3}$) on the value of p and its interaction with communication. The results show that being in the environment with a higher conflict of interests, makes the buyer less likely to take their ex-ante optimal action. In addition, we notice that communication has a differential effect across the two levels of conflict of interests: it significantly leads buyers away from their ex-ante optimal action when p=$\frac{1}{3}$, but has no significant effect when p=$\frac{2}{3}$. We can quantify the size of the communication effect in the p=$\frac{1}{3}$ case by computing the marginal effects from a probit regression of the buyer's decision being ex-ante optimal on a dummy for whether communication was present (and control variables - gender and age), using only the observations from the p=$\frac{1}{3}$ treatments. The results show that communication decreases the probability of taking the ex-ante optimal action in the p=$\frac{1}{3}$ case by 17.70 percentage points ($p - value = 0.054$).[4]

Table 7: Linear regression of buyer's decision being ex-ante optimal

	Dependent variable:
	Decision is ex-ante optimal
p=$\frac{1}{3}$	-0.240^{***}
	(0.083)
p=$\frac{1}{3}$ x CM	-0.176^{**}
	(0.082)
p=$\frac{2}{3}$ x CM	-0.061
	(0.082)
Female	-0.049
	(0.060)
Age	-0.013
	(0.009)
Constant	1.159^{***}
	(0.209)
Observations	232
R^2	0.138
Adjusted R^2	0.119
Residual Std. Error	0.442 (df = 226)
F Statistic	7.224^{***} (df = 5; 226)
Note:	*p<0.1; **p<0.05; ***p<0.01

When $p = \frac{2}{3}$, the observed deviations could be due to noise as their frequency is not particularly large. On the other hand, when $p = \frac{1}{3}$ the size of this frequency is larger both with and without communication. To explain this observation, we first need to notice that the ex-ante optimal action is "kind" towards the seller in the $p = \frac{2}{3}$ and "unkind" in the $p = \frac{1}{3}$ case. It is therefore possible that buyer's social preferences drive the deviations in the No_CM treatments. In the CM_ treatments communication might amplify the effect of social preferences leading buyers further

[4]See Appendix for probit regression table.

away from their ex-ante optimal action in the CM_13 treatment. Taken together, this evidence rules out Hypothesis 1.

Lastly, we can ask whether receivers can differentiate between truthful and untruthful sellers. To do so, we look at how often buyers choose to buy the red product after observing a message suggesting that the state is red, conditional on the true state. From figure 4 we note that truthful sellers are not more "persuasive" than untruthful ones as there is no significant difference between the frequencies of buying the red product after a truthful versus an untruthful message. Therefore, we can conclude that buyers are not able to separate truthful from untruthful sellers.

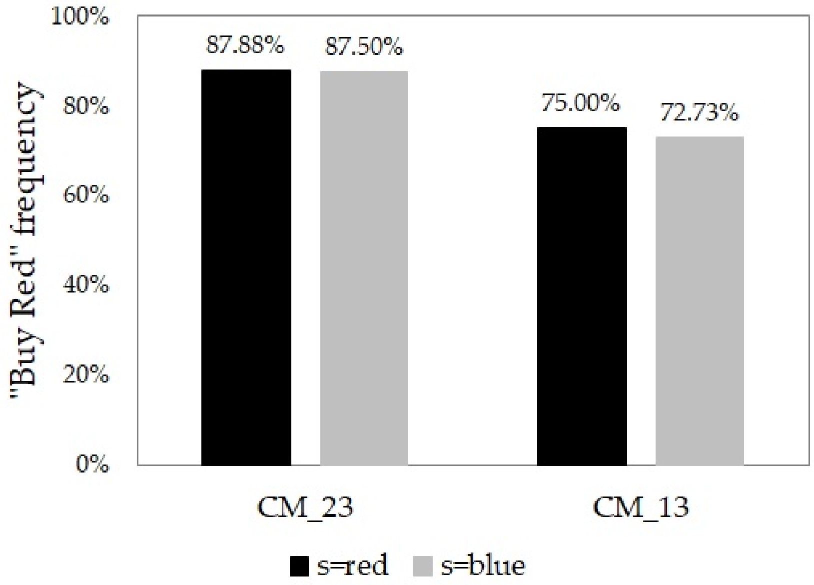


Figure 4: Frequency with which buyers choose $a = red$ after $m = red$ conditional on the underlying state

5.3 *Average payoffs*

In this section we report the effect of communication on players' average payoffs. Recall that if players have standard preferences, communication should have no effect on payoffs (Hypothesis 2). In the previous sections we have shown that players do not behave according to standard theoretical predictions. Nevertheless, the deviations are not high enough to render significant payoff differences according to our theoretical analysis from section 3.3. Table 8 presents the average payoffs comparison.

Table 8: Average payoffs across treatments

	$p=\frac{2}{3}$			$p=\frac{1}{3}$		
	CM_23	No_CM_23	p-value	CM_13	No_CM_13	p-value
Seller	0.793	0.862	0.528	0.567	0.393	0.063
Buyer	0.638	0.724	0.528	0.600	0.554	0.709

Note: p-values are based on a two-sided Barnard's test. Payoffs have been normalized to values of 1 and 0.

In line with the behavior analysis, we notice no significant payoff differences in the case of $p=\frac{2}{3}$, i.e. when interests are more likely aligned. However, when interests are more likely divergent, the seller's payoff is (weakly) significantly higher when messages are present. This is reflective of the higher proportion of buyers buying the red product in this treatment compared to the one where messages are absent. Result 5 follows:

Result 5. *When $p=\frac{2}{3}$, communication has no significant effect on players' average payoffs. When $p=\frac{1}{3}$, it significantly increases seller's average payoff and does not significantly impact buyer's average payoff.*

Result 5 contradicts Hypothesis 2. Furthermore, it goes against Hypothesis 3 since although we observe a positive share of truth-tellers, payoffs remain unaffected (at least from a statistical perspective). The result also contradicts Hypothesis 5 as when less than half of the sellers are lying averse, communication should have no effect in the CM_13 treatment. However, it supports Hypothesis 7 according to which there is an absolute effect of communication, across both levels of conflict of interests.

To make sure that Result 5 is not due to differences in the underlying distribution of states across treatments, we separate observations between those where a red ball was drawn and those where a blue ball was drawn. Table 9 presents this information.

Table 9: Average payoffs across treatments conditional on the underlying state

		CM_23	No_CM_23	p-value	CM_13	No_CM_13	p-value
s=red	Seller	0.838	0.905	0.530	0.727	0.421	0.058
	Buyer	0.838	0.905	0.530	0.727	0.421	0.058
s=blue	Seller	0.714	0.750	0.842	0.474	0.378	0.529
	Buyer	0.286	0.250	0.842	0.526	0.622	0.529

First, we note that there was no significant difference between the distribution of states in the CM_ treatments versus the No_CM treatments ($p - value = 0.528$ for $p = \frac{2}{3}$ and $p - value = 0.779$ for $p = \frac{1}{3}$, Barnard's two-sided test). Second, we find that when the state is blue, payoffs are not significantly different for either player irrespective of p. The significant positive effect of communication on the seller's payoff in the case of $p = \frac{1}{3}$ arises when the state is red ($p - value = 0.058$, Barnard's two-sided test). Hence, we can conclude that communication has a positive effect for sellers not because it enables them to deceive buyers (as in that case buyers' payoffs would be lower), but rather because it allows them to better signal the cases where interests are aligned (i.e. when the state is red).

How do these payoffs compare to what standard equilibrium theory would suggest? Table 10 presents the comparison of players' payoffs ("Obs.") with the standard equilibrium predictions ("Eq.").

Table 10: Average payoffs across treatments

	Seller's average payoff			Buyer's average payoff		
	Obs.	Eq.	p-value	Obs.	Eq.	p-value
CM_23	0.793	1	<0.001	0.638	0.638	1
CM_13	0.567	0	<0.001	0.600	0.633	0.792
No_CM_23	0.862	1	0.002	0.724	0.724	1
No_CM_13	0.393	0	<0.001	0.554	0.661	0.277

Note: p-values are based on a two-sided Barnard's test.

Result 6. *Buyers' average payoff is not significantly different than equilibrium predictions. Seller's average payoff is significantly lower than predicted when* $p = \frac{2}{3}$ *and significantly higher when* $p = \frac{1}{3}$.

Result 6 seems at odds with Result 5. To understand how these can be reconciled, we look at the distribution of buyers' errors over the state space. We define an error as the case where the buyer chooses $a = red$ when $s = blue$ or $a = blue$ when $s = red$. Table 11 presents the observed distribution of errors alongside the predicted one. The observed distribution of errors is significantly different than the predicted one for all treatments ($p - value < 0.010$ for all four treatments, Barnard's two-sided test).

From Table 11 we note that when $p = \frac{2}{3}$, the observed total number of errors is identical with the predicted one. However, buyers make a significantly higher number of errors when $s = red$. An error in this state is disadvantageous for the

Table 11: Distribution of errors

Treatment	$s = red$		$s = blue$	
	Observed	Predicted	Observed	Predicted
CM_23	6	0	15	21
CM_13	6	22	18	0
No_CM_23	4	0	12	16
No_CM_13	11	19	14	0

seller. On the other hand, when $s = blue$, buyers make fewer errors than predicted, meaning that they choose $a = blue$ more often than they should. This again is disadvantageous for the seller, but advantageous for the buyer. This is why the sellers' average payoff is significantly lower than predicted but the buyers' is not, as the increased number of errors buyers make when $s = red$ is compensated by the reduced number of errors when $s = blue$. A similar explanation holds for the case where $p = \frac{1}{3}$.

In addition, we observe that the distribution of errors across the state space is clearly skewed towards $s = blue$, irrespective of the treatment. This is another indication that buyers are not able to detect deception when communication is written and anonymous, a finding different than Belot and Van de Ven (2017) where communication was face-to-face but in line with the few experiments on deception detection with written communication (Chen and Houser, 2017).

6 CONCLUSION

In this paper I investigated if and when communication affects outcomes in a sender-receiver game with sender state-independent preferences and receiver state-dependent preferences. I hypothesized that the aversion of being profitably deceived (sucker aversion) can take over the positive effect that lying aversion might give rise to and lead to a negative effect of communication when the interests between the two parties are more likely to be conflicting. However, when a conflict of interests is relatively less likely, the opposite is the case. First, I found that communication does not harm when it was hypothesized to do so (i.e. when the likelihood of interests being aligned is relatively high). This is at odds with the results of Ert et al. (2014). A possible reason for the discrepancy in our results is the different complexity of environments used. The game implemented in Ert et al. (2014) is considerably more complex than the one I use. It could be that due to this, buyers in their experiment exert more cognitive effort to process the information. Under a higher cognitive load, psychological biases like sucker aversion are perhaps less likely to occur, and standard equilibrium play might have higher predictive power (Allred et al., 2016).

Second, the results in the current paper show that communication helps when the likelihood of interests being aligned is relatively low by increasing sellers' payoff without reducing buyers' payoff. Sellers benefit from the fact that buyers trust messages more often than they ought to. At the same time, buyers are not affected by their gullibility since some sellers are lying averse, reporting blue states truthfully more often than standard equilibrium theory would suggest. Nevertheless, this is not predicted by the theoretical analysis since the observed amount of truth-telling is much lower than the theoretically necessary level. An explanation for this finding could be that buyers overweight the probability of interacting with a lying averse seller. This is in line with Sheremeta and Shields (2017) who find that in a setting similar to the one analyzed in this paper but where the binary states are uniformly distributed, receivers are too optimistic about senders' truthful behavior. Importantly, participants in their experiment play simultaneously the role of sender and that of receiver. Hence, beliefs about senders' behavior can be confounded with self-image concerns.

Third, I find a surprising deviation from standard equilibrium predictions in the treatments where communication is absent. In the case where interests are more likely to be aligned, buyers choose the action that harms the sellers significantly more often than predicted. This is an indication that something different than sucker aversion is at play. An explanation based on probability matching (Myers, 2014; Vulkan, 2000) is ruled out as subjects do not make repeated choices. I propose two tentative explanations, none of which can capture the entire pattern of behavior. One is that buyers have different levels of strategic sophistication and those with a lower level display a higher degree of skepticism. This though is contradicting the effect observed in the case where interests alignment is relatively less likely. In this treatment, buyers choose the action preferred by the seller significantly more often. The second possibility is that social preferences matter, making buyers more likely to choose the action that is advantageous for the seller irrespective of the level of conflict. This would explain the latter result, but not the former according to which communication has a differential effect across the two levels of conflict of interests.

I note, lastly, that in the treatments where interests are more likely to be aligned, the effect of social preferences goes in the opposite direction with that of communication under sucker aversion. In the treatments where interests - alignment is less likely, these effects would push behavior in the same direction. Therefore, a possibility is that the effect of social preferences and of sucker aversion cancel out in the former case, while communication enhances social preferences in the latter case. Further research on the interplay between these two behavioral drivers is welcomed.

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

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

Allred, S., Duffy, S., and Smith, J. (2016). Cognitive load and strategic sophistication. *Journal of Economic Behavior & Organization*, 125:162–178.

Belot, M., Bhaskar, V., and Van De Ven, J. (2012). Can observers predict trustworthiness? *Review of Economics and Statistics*, 94(1):246–259.

Belot, M. and Van de Ven, J. (2017). How private is private information? the ability to spot deception in an economic game. *Experimental Economics*, 20(1):19–43.

Bicchieri, C. and Lev-On, A. (2007). Computer-mediated communication and cooperation in social dilemmas: An experimental analysis. *Politics, Philosophy & Economics*, 6(2):139–168.

Blume, A., DeJong, D. V., Kim, Y.-G., and Sprinkle, G. B. (1998). Experimental evidence on the evolution of meaning of messages in sender-receiver games. *The American Economic Review*, 88(5):1323–1340.

Blume, A., DeJong, D. V., Kim, Y.-G., and Sprinkle, G. B. (2001). Evolution of communication with partial common interest. *Games and Economic Behavior*, 37(1):79–120.

Brosig, J. (2002). Identifying cooperative behavior: some experimental results in a prisoner's dilemma game. *Journal of Economic Behavior & Organization*, 47(3):275–290.

Cabrales, A., Feri, F., Gottardi, P., and Meléndez-Jiménez, M. A. (2020). Can there be a market for cheap-talk information? An experimental investigation. *Games and Economic Behavior*, 121:368–381.

Cai, H. and Wang, J. T.-Y. (2006). Overcommunication in strategic information transmission games. *Games and Economic Behavior*, 56(1):7–36.

Capraro, V. (2018). Gender differences in lying in sender-receiver games: A meta-analysis. *Judgment and Decision Making*, 13(4):345–355.

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

Charness, G. and Dufwenberg, M. (2011). Participation. *American Economic Review*, 101(4):1211–37.

Chen, J. and Houser, D. (2017). Promises and lies: can observers detect deception in written messages. *Experimental Economics*, 20(2):396–419.

Cooper, R., DeJong, D. V., Forsythe, R., and Ross, T. W. (1992). Communication in coordination games. *The Quarterly Journal of Economics*, 107(2):739–771.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–1451.

Darke, P. R. and Ritchie, R. J. (2007). The defensive consumer: Advertising deception, defensive processing, and distrust. *Journal of Marketing Research*, 44(1):114–127.

Dawes, R. M., McTavish, J., and Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, 35(1):1.

Dickhaut, J., Ledyard, M., Mukherji, A., and Sapra, H. (2003). Information management and valuation: an experimental investigation. *Games and Economic Behavior*, 44(1):26–53.

Dickhaut, J. W., McCabe, K. A., and Mukherji, A. (1995). An experimental study of strategic information transmission. *Economic Theory*, 6(3):389–403.

Ert, E., Creary, S., and Bazerman, M. (2014). Cynicism in negotiation: When communication increases buyers' skepticism. *Judgment and Decision Making*, 9:191–198.

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

Forsythe, R., Lundholm, R., and Rietz, T. (1999). Cheap talk, fraud, and adverse selection in financial markets: Some experimental evidence. *The Review of Financial Studies*, 12(3):481–518.

Friestad, M. and Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21(1):1–31.

Gneezy, U., Kajackaite, A., and Sobel, J. (2018). Lying aversion and the size of the lie. *American Economic Review*, 108(2):419–53.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Kawagoe, T. and Takizawa, H. (2009). Equilibrium refinement vs. level-k analysis: An experimental study of cheap-talk games with private information. *Games and Economic Behavior*, 66(1):238–255.

Khalmetski, K., Rockenbach, B., and Werner, P. (2017). Evasive lying in strategic communication. *Journal of Public Economics*, 156:59–72.

Lipnowski, E. and Ravid, D. (2017). Cheap Talk with Transparent Motives. *Available at SSRN: https://ssrn.com/abstract=2941601 or http://dx.doi.org/10.2139/ssrn.2941601*.

Minozzi, W. and Woon, J. (2013). Lying aversion, lobbying, and context in a strategic communication experiment. *Journal of Theoretical Politics*, 25(3):309–337.

Mohlin, E. and Johannesson, M. (2008). Communication: Content or relationship? *Journal of Economic Behavior & Organization*, 65(3-4):409–419.

Myers, J. L. (2014). Probability learning and sequence learning. *Handbook of Learning and Cognitive Processes, ed. WK Estes*, pages 171–205.

Ockenfels, A. and Selten, R. (2000). An experiment on the hypothesis of involuntary truth-signalling in bargaining. *Games and Economic Behavior*, 33(1):90–116.

Sánchez-Pagés, S. and 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. and Vorsatz, M. (2009). Enjoy the silence: an experiment on truth-telling. *Experimental Economics*, 12(2):220–241.

Serra-Garcia, M., van Damme, E., and Potters, J. (2011). Hiding an inconvenient truth: Lies and vagueness. *Games and Economic Behavior*, 73(1):244–261.

Sheremeta, R. M. and Shields, T. W. (2017). Deception and reception: The behavior of information providers and users. *Journal of Economic Behavior & Organization*, 137:445–456.

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, 14(1):101–118.

APPENDIX

A PROOFS

Lemma 3

Proof. Let the L seller's strategy be $(m_L(R), m_L(B)) = (R, R)$. Buyer's beliefs consistent with the specified seller strategy are:

$$\mu(R) = \frac{pr(m = R|s = R) \cdot pr(s = R)}{pr(m = R)} = \frac{1 \cdot p}{p + (1-p) \cdot \lambda} = \frac{p}{p + (1-p) \cdot \lambda}$$

$$\mu(B) = \frac{pr(m = B|s = R) \cdot pr(s = R)}{pr(m = B)} = 0$$

Given these beliefs, the buyer best replies by maximizing his expected utility given each message.

$$EU_{buyer}(a = R, m = R) = 1 \cdot \mu(R) + 0 \cdot (1 - \mu(R)) = \frac{p}{p + (1-p) \cdot \lambda}$$

$$EU_{buyer}(a = B, m = R) = 0 \cdot \mu(R) + 1 \cdot (1 - \mu(R)) = \frac{(1-p) \cdot \lambda}{p + (1-p) \cdot \lambda}$$

$$EU_{buyer}(a = R, m = B) = 1 \cdot \mu(B) + 0 \cdot (1 - \mu(B)) = 0$$

$$EU_{buyer}(a = B, m = B) = 0 \cdot \mu(B) + 1 \cdot (1 - \mu(B)) = 1$$

This means that $EU_{buyer}(a = B, m = B) > EU_{buyer}(a = R, m = B)$.

When $p \geq \frac{1}{2}$, since $\lambda < 1$ we have that $EU_{buyer}(a = R, m = R) > EU_{buyer}(a = B, m = R)$. Therefore, the buyer's best reply strategy is $(a(R), a(B)) = (R, B)$. Does any L seller type have an incentive to deviate given the buyer's strategy? Clearly, when $s = R$, no L seller type would like to deviate by sending $m_L(R) = B$ since that would lead to a payoff of 0, which is smaller than the payoff the seller gets from $m_L(R) = R$. When $s = B$, $EU_L(m = R, s = B) = 1$ while $EU_L(m = B, s = B) = 0$, so the L seller does not want to deviate from the original strategy $(m_L(R), m_L(B)) = (R, R)$.

When $p < \frac{1}{2}$ the buyer best replies by choosing $(a(R), a(B)) = (R, B)$ when $\lambda \leq \frac{p}{1-p}$ and $(a(R), a(B)) = (B, B)$ when $\lambda > \frac{p}{1-p}$. The L seller has no incentive to deviate in either case implying that the specified strategies constitute a PBE of the L-game where $p < \frac{1}{2}$.

□

Lemma 4

Proof. Let the L seller's strategy be $(m_L(R), m_L(B)) = (B, B)$. Buyer's beliefs consistent with this strategy are: $\mu(R) = 1$, $\mu(B) = \frac{p\lambda}{1-p+p\lambda}$. Consequently, his expected utilities from either choice are: $EU_{buyer}(a = R, m = R) = 1$, $EU_{buyer}(a = B, m = R) = 0$, $EU_{buyer}(a = R, m = B) = \frac{p\lambda}{1-p+p\lambda}$, $EU_{buyer}(a = B, m = B) = \frac{1-p}{1-p+p\lambda}$.

Therefore, the buyer best replies by choosing $a(R) = R$ irrespective of λ.

When $m = B$, his optimal choice depends on p and λ as follows. The buyer best replies by choosing $a(B) = B$ if $\lambda < \frac{1-p}{p}$. However, this cannot be supported in equilibrium since the L seller can profitably deviate to $(m_L(R), m_L(B)) = (R, R)$.

The buyer best replies by choosing $a(B) = R$ if $\lambda \geq \frac{1-p}{p}$. This condition is satisfied only if $p \geq \frac{1}{2}$. If $p < \frac{1}{2}$, the condition would require that λ is greater than 1 which is not possible.

Hence, if $p \geq \frac{1}{2}$ and $\lambda > \frac{1-p}{p}$, then the receiver's best reply is $a(B) = R$. The L seller has no profitable deviation since the buyer's action is always R irrespective of the message.

If $p < \frac{1}{2}$, then $(m_L(R), m_L(B)) = (B, B)$ cannot be part of an equilibrium of the L-game.

□

Lemma 5

Proof. Let the L seller's strategy be $(m_L(R), m_L(B)) = (R, B)$. Buyer's beliefs consistent with this strategy are: $\mu(R) = 1$, $\mu(B) = 0$. Consequently, his expected utilities from either choice are: $EU_{buyer}(a = R, m = R) = 1$, $EU_{buyer}(a = B, m = R) = 0$, $EU_{buyer}(a = R, m = B) = 0$, $EU_{buyer}(a = B, m = B) = 1$. Therefore, the buyer's best reply is: $(a(R), a(B)) = (R, B)$.

Given the buyer's best reply, the L seller has a profitable deviation from $(m_L(R), m_L(B)) = (R, B)$ to $(m_L(R), m_L(B)) = (R, R)$ since that ensures the seller an expected payoff of 1 rather than $p \leq 1$.

Therefore, the $(m_L(R), m_L(B)) = (R, B)$, cannot be supported in a PBE irrespective of the value of p.

□

Lemma 6

Proof. Let the L seller's strategy be $(m_L(R), m_L(B)) = (B, R)$. Buyer's beliefs consistent with this strategy are: $\mu(R) = \frac{(1-\lambda)p}{p+\lambda-2p\lambda}$, $\mu(B) = \frac{p\lambda}{1-p-\lambda+2p\lambda}$.

Consequently, his expected utilities from either choice are: $EU_{buyer}(a = R, m = R) = \frac{p-p\lambda}{p+\lambda-2p\lambda}$, $EU_{buyer}(a = B, m = R) = \frac{\lambda-p\lambda}{p+\lambda-2p\lambda}$, $EU_{buyer}(a = R, m = B) = \frac{p\lambda}{1-p-\lambda+2p\lambda}$, $EU_{buyer}(a = B, m = B) = \frac{1-p-\lambda+p\lambda}{1-p-\lambda+2p\lambda}$.

If $p \geq \frac{1}{2}$ then $p + \lambda - 2p\lambda > 0$ and $1 - p - \lambda + 2p\lambda > 0$. In this case, if $\lambda < p < 1 - \lambda$, the buyer best replies by choosing $(a(R), a(B)) = (R, B)$. However, in this case the L seller can profitably deviate by choosing $(m_L(R), m_L(B)) = (R, R)$ and so, these strategies cannot represent an equilibrium. A similar argument is valid for when $\lambda > p > 1 - \lambda$, except that receiver's best reply is $(a(R), a(B)) = (B, R)$ in this case.

However, if $p > \lambda > 1 - p$, the buyer's best reply is $(a(R), a(B)) = (R, R)$. Now, the L seller has no incentive to deviate from $(m_L(R), m_L(B)) = (B, R)$ and these strategies represent a PBE of the L-game where $p \geq \frac{1}{2}$. Similarly, if $\lambda > p > 1 - p$, the buyer's best reply is $(a(R), a(B)) = (B, B)$ and the seller has no incentive to deviate.

If $p < \frac{1}{2}$ then $p + \lambda - 2p\lambda < 0$ and $1 - p - \lambda + 2p\lambda > 0$. In this case, the seller's strategy cannot be supported in a PBE whenever the buyer's best reply is a different action for each message. Specifically, when $p < \lambda < 1 - p$, the buyer's best reply is $(a(R), a(B)) = (R, B)$. In this case, the seller has a profitable deviation to $(m_L(R), m_L(B)) = (R, R)$. $p > \lambda > 1 - p$ is not possible when $p < \frac{1}{2}$.

However, if $\lambda > 1 - p$, leading to a buyer's best reply of $(a(R), a(B)) = (R, R)$ or, if $\lambda < p$, leading to a buyer's best reply of $(a(R), a(B)) = (B, B)$, the L seller has no incentive to deviate from $(m_L(R), m_L(B)) = (B, R)$ and these strategies represent a PBE of the L-game where $p < \frac{1}{2}$.

□

Lemma 7

Proof. Since the K buyer's action is fixed, the only thing we need to consider when asking whether the N buyer would like to deviate from the equilibrium described by Lemma 1 is whether the proportion of K buyers influences the optimal action for the N buyer. This would happen only if the seller would be incentivized to transmit more persuasive information in equilibrium the higher p_k is.

Suppose $p_k = 1$. In this case, the seller is indifferent between telling the truth or any other strategy. However, as long as $p_k < 1$, the seller would strictly prefer to send his ex ante optimal message (non-persuasive; see Lemma 1) irrespective of the state, as this makes it optimal for the N buyer to choose his ex ante optimal strategy $(a_N(R), a_N(B)) = (R, R)$. This way, the seller maximizes his expected payoff.

□

B ADDITIONAL TABLES

Table 12: Marginal effects from probit regression of buyer's decision being optimal (p=$\frac{1}{3}$ only)

	Dependent variable:
	Decision is ex-ante optimal
CM	-0.177^{*}
	(0.092)
Female	-0.089
	(0.102)
Age	-0.014
	(0.012)
Observations	116
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 13: Marginal effects from probit regression of buyer's decision being optimal (p=$\frac{2}{3}$ only)

	Dependent variable:
	Decision is ex-ante optimal
CM	-0.061
	(0.092)
Female	-0.012
	(0.102)
Age	-0.015
	(0.012)
Observations	116
Note:	*p<0.1; **p<0.05; ***p<0.01

C INSTRUCTIONS

Welcome and thank you for participating in this experiment. Throughout the whole experiment you must remain seated and mobile phones and other electronic devices must be switched off. If there are any questions please raise your hand and an experimenter will come to answer your questions in private.

Payment: In this experiment you can earn points. At the end of the experiment you will be paid **£3 as a participation fee** plus **£0.10 for each point earned**. You will be paid in private and in cash.

All your decisions are anonymous, so your identity will be kept secret at all times.

Brief description of the task:

At the beginning of the experiment you will be randomly matched with another participant. One of you will have the role of Buyer and the other the role of Seller.

There are two products: a Red product and a Blue product. The Buyer must buy one of these products. The Seller knows the qualities of the two products but the Buyer does not. The final payoff to both participants depends on the buying choice and the quality of the chosen product. A more detailed description of the task is given below.

Detailed description of the task:

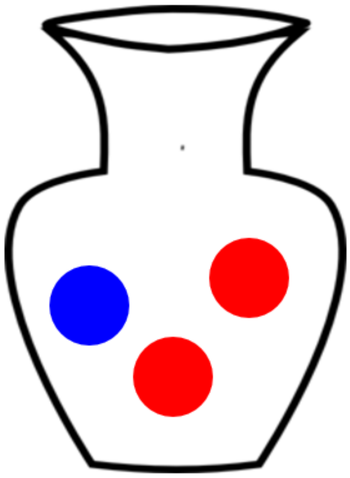
The task consists of 2 stages.

Stage 1: *The Seller is informed about the qualities of the two products [and sends a message]*[5].

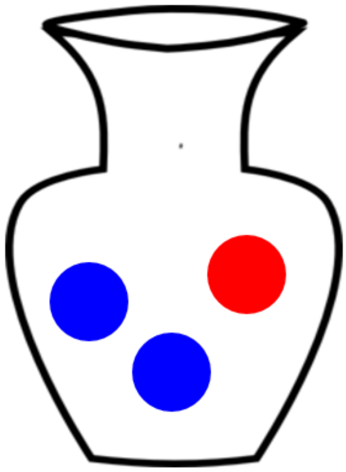
In this stage, the computer will randomly draw a ball from a computerized urn containing 3 balls: 2 red and 1 blue. The urn is depicted below:

[5] The content within square brakets was present only in the CM_{-} treatments

(a) Urn for _23 treatments



(b) Urn for _13 treatments



The Seller will observe the colour of the drawn ball, but the Buyer will not.

- If the colour of the drawn ball is **red**, the **Red product** is **better than** the **Blue product**. The Red product is worth 30 points to the Buyer and the Blue product is worth nothing to the Buyer.
- If the colour of the drawn ball is **blue**, the **Blue product** is **better than** the **Red product**. The Blue product is worth 30 points to the Buyer and the Red product is worth nothing to the Buyer.

[After observing the qualities of the two products, the Seller will send a message to the Buyer, by typing in a chat box. There is no restriction regarding the content of the message, except that Sellers are not allowed to reveal any information that might identify them (e.g. name, seat number, etc.)]

Stage 2: The Buyer [observes the message and] decides whether to buy the Red product or the Blue product.

After the Seller observes the qualities of the products [and after observing the message from the Seller], the Buyer will decide between buying either the Red product or the Blue product.

How your earnings are determined:

The Buyer will earn 30 points if they buy the better product. The Seller will earn 30 points if the Buyer buys the Red product. Thus, the point-earnings will be determined as follows.

The Buyer

- earns 30 points if a Red ball is drawn and she buys the Red product;
- earns nothing if a Red ball is drawn and she buys the Blue product;
- earns 30 points if a Blue ball was drawn and she buys the Blue product;
- earns nothing if a Blue ball was drawn and she buys the Red product.

The Seller

- earns 30 points if the Buyer buys the Red product;
- earns nothing if the Buyer buys the Blue product.

This is summarised in the Table below:

	Buy the Red product	**Buy the Blue product**
Red ball drawn	The Seller earns 30 points, The Buyer earns 30 points	The Seller earns 0 points, The Buyer earns 0 points
Blue ball drawn	The Seller earns 30 points, The Buyer earns 0 points	The Seller earns 0 points, The Buyer earns 30 points

Preliminary questions: Before the experiment begins, you will be asked to answer a few questions regarding your understanding of the instructions. The experiment will begin only after all participants have answered these questions correctly.

Final questionnaire: After everyone makes their decisions, you will be asked to fill in a short questionnaire. You will then be paid your earnings in private and in cash. You will receive a £3 participation fee plus £0.10 for each point that you earn.