Why do promises affect trustworthiness, or do they?

Ismayilov, Huseyn; Potters, Jan

Published in:
Experimental Economics

DOI:
10.1007/s10683-015-9444-1

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Exp Econ (2016) 19:382–393
DOI 10.1007/s10683-015-9444-1

ORIGINAL PAPER

Why do promises affect trustworthiness, or do they?

Huseyn Ismayilov[1,2] · **Jan Potters**[2]

Received: 28 June 2013 / Revised: 2 May 2015 / Accepted: 10 May 2015 /
Published online: 20 May 2015

Abstract We set out to test whether the effect of promises on trustworthiness derives from the fact that they are made (internal consistency) or that they are received (social obligation). The results of an experimental trust game appeared at first to support the former mechanism. Even when trustee messages are not delivered to trustors, trustees who make a promise are more likely to act trustworthy than those who do not make a promise. However, we subsequently ran a control treatment with restricted (non-promise) communication to examine whether the correlation between promises and trustworthiness is causal. The results show that the absence of promises does not decrease average cooperation rates. This indicates that promises do not induce trustworthiness, they are just more likely to be sent by cooperators than by non-cooperators.

Keywords Promises · Communication · Commitment · Endogeneity · Experiment

Mathematics Subject Classification C91 · D03 · D82 · L15

Electronic supplementary material The online version of this article (doi:10.1007/s10683-015-9444-1) contains supplementary material, which is available to authorized users.

✉ Huseyn Ismayilov
hismayilov@ada.edu.az

Jan Potters
j.j.m.Potters@uvt.nl

[1] The School of Business, ADA University, 11 Ahmadbey Aghaoglu Street, 1008 Baku, Azerbaijan

[2] CentER and Netspar, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

1 Introduction

Communication is often found to foster trust and cooperation and many studies have emphasized the role of promises in this respect (Belot et al. 2010; Vanberg 2008; Bicchieri and Lev-on 2007; Charness and Dufwenberg 2006; Sally 1995; Ostrom et al. 1992; Orbell et al. 1988). A prime explanation for the impact of promises is that they create a commitment. In the present paper we set out to explore the nature and force of this commitment. Specifically, we examine whether a promise has commitment power because the promisor makes it or because the promisee learns about it.

A preference for promise-keeping may derive from a more general preference for consistency (see Ellingsen and Johannesson 2004; Kerr and Kaufman-Gilliland 1994). If a person has expressed that she will do X, not doing X creates an inconsistency which the person may want to avoid.[1] Whether or not a person's statement (promise) is consistent with the person's action does not depend on whether someone else may be affected by the statement or even learns about it. What counts for the individual is that she has expressed an intention to do something; as a consequence she prefers to take an action which matches that intention. We call this the *internal consistency explanation* for promise keeping. An alternative interpretation of the commitment-based explanation for promise-keeping is that people feel obliged to fulfill verbal contracts and agreements (Vanberg 2008). This conceptualization of the commitment requires, not only that the promisor made the promise, but also that the promisee learns about it. We call this the *social obligation explanation* for promise keeping.

To assess the empirical support for these alternative forms of commitment, we tweaked the experimental trust game by Charness and Dufwenberg (2006). Trustees had an opportunity to write a pre-play free-form message to trustors, but the message was delivered to the trustor only with probability $\frac{1}{2}$.[2] Thus, in our experiment 50 % of the trustees wrote a message that was not delivered to the trustors.

The results show that trustees who made a promise were significantly more likely to act trustworthy than trustees who did not make a promise, irrespective of whether the message was delivered or not. Conditional on messages being delivered, promisors were 12 % more likely to act trustworthy than non-promisors (54 vs. 42 %); conditional on messages not being delivered, promisors were 21 % more likely to act trustworthy than non-promisors (35 vs. 14 %). These results are in line with the internal consistency hypothesis that promises create commitment even when not delivered.

An important caveat, of course, is that promises are endogenous.[3] It may be that trustworthy trustees are more likely to make a promise than untrustworthy trustees, in which case self-selection may drive the difference between promisors and non-promisors. To examine whether promises increase trustworthiness, we later ran a

[1] A preference for consistency is also in line with an aversion towards lies (see, e.g., Lopez-Perez and Spiegelman 2013; Lundquist et al. 2009). Serra-Garcia et al. (2013) suggest, however, the preference for promise-keeping is even stronger than the preference for truth-telling.

[2] Our randomization of message delivery is similar to Vanberg (2008) random replacement of partners.

[3] We thank the reviewers for emphasizing this point.

control treatment in which trustees could send messages but were not allowed to make promises. It turned out that in the control treatment trustees were not less trustworthy than they were in the base treatment in which they could and did make promises; actually they were somewhat, insignificantly, more trustworthy (49 % in the control treatment vs. 39 % in our base treatment). This suggests that the correlation between promises and trustworthiness is due to self-selection. Promises do not increase trustworthiness in our experiment; they are just more likely to be sent by cooperators than by non-cooperators.

The fact that promises are endogenous is not unique to our experiment. Many studies which report an impact of promises on cooperation face a potential problem of reverse causality. The fact that a treatment with communication leads to more cooperation than a treatment without communication does not by itself help to draw causal inferences on the effect of promises. Communication can have an effect *per se*, just as message delivery increases trustworthiness in our base treatment. There we find that promises were more likely to be kept if they were delivered (54 %) than if they were not (35 %), but trustees who did not make a promise were also more likely to be trustworthy if their messages were delivered (42 %) than if they were not (14 %). This suggests that the positive impact of communication on trustworthiness need not be related to promises, even if there is a positive correlation between promises and trustworthiness.

2 Experimental design and procedure

2.1 Experimental design

To generate observations of trustee messages which can affect trustors, as well as messages which cannot, we introduce a twist to the trust game employed by Charness and Dufwenberg (2006). This trust game is described in Fig. 1. There are two players in this game, A and B. First, A chooses to play *In* or *Out*. Next, B chooses *Roll* or *Don't Roll* a six sided die. If A chooses *Out*, then B's choice is irrelevant and both players get 5 Euros. If A chooses *In* and B chooses *Don't Roll*, A receives 0 and B receives 14 Euros. Finally, if A plays *In* and B plays *Roll*, then B gets 10 Euros and rolls a six sided die to determine the payoff to A. A receives 12 Euros with probability $\frac{5}{6}$ and 0 with probability $\frac{1}{6}$.

As in Charness and Dufwenberg (2006), we allow B to write a pre-play message to A. However, in our design with probability $\frac{1}{2}$ a message sheet is not delivered to A. This is known to both A and B. After writing a message, B learns whether his message sheet will be delivered to A or not from the outcome of a random draw. If A receives an empty sheet, A knows that the message sheet by B was not chosen to be delivered. The timeline for the pre-play message stage is shown in Fig. 2. After the pre-play message stage, the trust game depicted above is played. Instructions are provided in Online Appendix A.

With this design, we obtain observations of messages from B which are delivered to A and observations of messages from B which are not delivered to A. In what

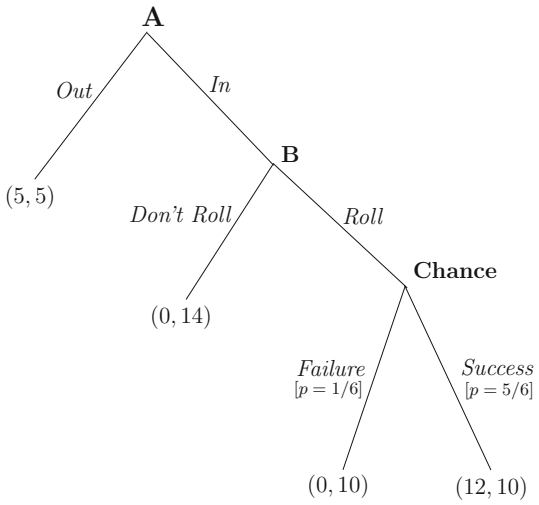

Fig. 1 Trust game of Charness and Dufwenberg (2006)

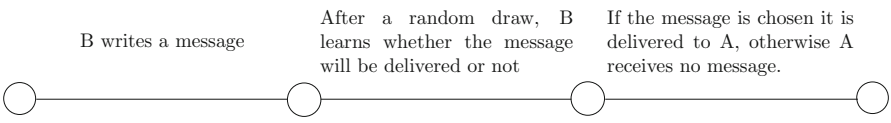

Fig. 2 Timeline of the pre-play message stage

follows we call the former the *Message delivered* condition and the latter the *Message not delivered* condition. Within each condition there will be some Bs who make a promise to *Roll* and some who do not make a promise to *Roll*.

Several experiments have shown that when given the opportunity to send pre-play messages subjects who send a promise to be trustworthy are more likely to cooperate than subjects who do not send such a promise (see, for example, Charness and Dufwenberg 2006; Ellingsen and Johannesson 2004; Vanberg 2008). Our aim is to test whether this effect of promises derives from the fact that they are made or from the fact that they are received.

Hypothesis 1 (*internal consistency*): B players who make a promise to *Roll* are more likely to *Roll* than B players who do not make a promise, irrespective of whether the message is delivered or not.

Hypothesis 2 (*social obligation*): B players who make a promise to *Roll* are more likely to *Roll* than B players who do not make a promise if and only if the message is delivered.

2.2 Experimental procedure

The experiment was conducted at the CenterLab, Tilburg University. Subjects were students recruited via email invitations. 12 sessions were conducted with a total of 260 participants (there were 20 subjects per session in 7 sessions, and 24 subjects per

session in 5 sessions). Average earnings were around 11 Euros per session (including a 3 Euros show-up fee). The duration of each session was approximately one hour.

Subjects were seated behind visually partitioned workstations. At the beginning the instructions were distributed and read aloud. Questions were answered privately. Half of the subjects were assigned the role of A and the other half the role of B. Each A was matched with a B to form a pair. Sheets with identification numbers and a letter B on top were distributed to all Bs. Each B knew his or her identification number, but no other subject did. We allowed enough time for all Bs to write a message to A in his or her pair. If B did not want to write a message he or she could circle the letter B on top of the sheet. After all Bs finished writing a message and put their message sheets face down, the experimenter collected all message sheets. The experimenter quickly checked the compliance of the messages with anonymity rules. Then, the identification numbers of all Bs were shuffled and exactly half of them were randomly chosen and publicly revealed. The message sheets of those Bs whose numbers were chosen were distributed to the respective As. Before the message sheets were delivered to the As, the identification numbers of the Bs were publicly cut off so that all players would be further assured that the messages were anonymous. The message sheets of Bs whose numbers were not chosen were replaced by empty sheets. Thus, in all pairs A received a sheet, either empty or with a message, depending on whether a message sheet was chosen to be delivered in that pair or not. Note that an empty sheet was different from a delivered message sheet without text, since the latter had the letter B circled on top. The identity of subjects in pairs was not revealed at any time.

After the message sheets were delivered to the respective As, the game depicted in Fig. 1 was played using the strategy method. That is B chose *Roll* or *Don't Roll* before knowing A's choice for *In* or *Out*. Unlike the pre-play message stage, the actual game stage was computerized using the Z-tree software (Fischbacher 2007). Subjects entered choices on their screens. After choices were made by all As and Bs the experimenter approached each B to roll a die. To ensure anonymity all Bs rolled a die irrespective of their choice and entered the outcome of the die roll on their screen. The game was played for one round only. After the payoffs were realized subjects were paid privately and in cash.

Finally, we elicit subjects' expectations to control for beliefs and to test the predictions under the expectations based guilt aversion explanation for promise keeping suggested by Charness and Dufwenberg (2006). We refer interested readers to Charness and Dufwenberg (2006) for details of this explanation. We followed Vanberg (2008) in revealing beliefs of players with some differences. For details see Online Appendix B.

3 Results

In total we obtained observations for 130 pairs, 65 pairs each in the *Message not delivered* condition and in the *Message delivered* condition. We hired three research assistants to code each message as a promise or no promise (Coder Instructions are provided in Online Appendix D). For our analysis we classified messages based on

Table 1 Promises and *Roll* rates[a]

Condition	B's *Roll* rate			
	Promise	No promise	Z stat	Row total
Message not delivered	15/43	3/22	1.81**	18/65
	(35 %)	(14 %)		(28 %)
Message delivered	25/46	8/19	0.90	33/65
	(54 %)	(42 %)		(51 %)
Z stat	1.84**	2.05**	–	2.69***
Column total	40/89	11/41	1.97**	51/130
	(45 %)	(27 %)		(39 %)

[a] The Z stat reflects two sample proportion test for the two populations

*, **, and *** denote significance at $p \leq 0.10$, $p \leq 0.05$, and $p \leq 0.01$ respectively for one tailed test

the majority decision by coders (109 out of 130 decisions were unanimous). The classification is available in Online Appendix C. For both conditions combined, 89 out of 130 Bs (68 %) made a promise to *Roll*: 43 out of 65 Bs (66 %) in the *Message not delivered* condition and 46 out of 65 Bs (71 %) in the *Message delivered* condition.

Table 1 presents *Roll* rates by Bs who made a promise and by Bs who did not for each condition separately and for the combined data. For the combined data, the *Roll* rates are higher for those who made a promise (45 %) than for those who did not (27 %) and this difference is statistically significant. In the *Message not delivered* condition the *Roll* rates by Bs are significantly higher for those who made a promise (35 %) than for those who did not (14 %). In the *Message delivered* condition the difference in *Roll* rates is 12 %, but it is statistically insignificant. This difference is smaller than that obtained for the *Messages* (5,5) treatment in Charness and Dufwenberg (2006) (12 vs. 19 %). These results are in line with Hypothesis 1 (internal consistency) which postulates that promises increase trustworthiness both when messages are delivered and when messages are not delivered.

The results also show that for trustees who made a promise the *Roll* rates were significantly higher if a promise was delivered to the trustor than if it was not (54 vs. 35 %). Clearly, a delivered promise was more likely to be kept than an undelivered promise. Note, however, that there was also a positive effect of messages being delivered on *Roll* rates of trustees who did not make a promise (42 vs. 14 %). In other words, we observe increased *Roll* rates with message delivery

[4] One concern one may have is that messages which are coded as not being a promise are actually intended and perceived to be a promise by the sender. However, if we only consider blank messages, we observe a significant increase in *Roll* rates due to message delivery (4 out of 7, 57 %, when delivered compared to 0 out of 10, 0 %, when not delivered, $p = 0.015$, two-tailed Fisher exact test). Somewhat surprisingly, if we look only at non-blank non-promise messages delivery has a somewhat weaker effect on *Roll* rates (4 out of 12, 33 %, when messages are delivered and 3 out of 12, 25 %, when messages are not delivered, Z stat = 0.45, $p = 0.65$, two-tailed proportions test).

(communication) not only for trustees who made a promise but also for trustees who did not make a promise.[4]

In Online Appendix B (Table B.4) we report the estimates from linear probability model regressions which show a significant effect of promises and message delivery on *Roll* rates, but an insignificant effect of the interaction effect between the two. We also show that these results are hardly affected when we control for the second-order beliefs of the trustees. The latter result means that the correlation between promises and *Roll* rates in the *Messages not delivered* condition and the effect of messages being delivered on nonpromisors cannot be explained by changes in the second-order beliefs of the trustees.

At first glance, our results support Hypothesis 1 (internal consistency). The fact that in the *Message not delivered* condition subjects who made a promise were more likely to *Roll* than subjects who did not make a promise suggests that making a promise creates a commitment even when it is not delivered.[5] However, this correlation is prone to an endogeneity problem. Specifically, it is possible that those who make a promise are more likely to be trustworthy not because they are affected by a promise made but because trustworthy trustees are more likely to make a promise than untrustworthy ones.

To clarify this issue, we subsequently ran a new treatment with restricted communication in which B players were not allowed to send messages that refer to the game being played (hence, B players could not make a promise). In all other respects, this treatment was identical to our original treatment with unrestricted communication. We call this new treatment *Restricted Communication* and our original treatment *Unrestricted Communication*. If promises, delivered or not, cause trustworthiness we should expect higher *Roll* rates in the *Unrestricted communication treatment* than in the *Restricted communication treatment* because promises can be made in the former but not in the latter treatment.

Table 2 reports *Roll* rates by condition both for the *Restricted communication* treatment and for the *Unrestricted communication* treatment. The absence of promises in the *Restricted Communication* treatment does not result in lower *Roll* rates both when messages are delivered and when they are not delivered.[6] This suggests that promises, whether delivered or undelivered, do not create

[5] It might be argued that when messages were not delivered promises might be correlated with trustworthiness not because of a cost of breaking a promise per se but because the messages were observed by the experimenter. While our experimental procedures were not double blind, it was practically impossible for the experimenter to remember all messages sent by trustees and then map them to individuals and choices. Note that the messages were handwritten while the choices for the trust game were entered on the computer screen. This was made clear to subjects in instructions. Additional evidence is provided by Deck et al. (2013). The authors run a single-blind and a double-blind protocols of the trust game with pre-play messages of Charness and Dufwenberg (2006) and find no difference between the two protocols. This result suggests that 'an experimenter effect' is not an issue in the trust game with pre-play messages.

[6] The proportion of blank messages was 13 % in the *Unrestricted Communication* as opposed to 24 % in the *Restricted Communication* treatment. The difference is marginally significant at $p = 0.083$ with a two-tailed proportions test. Although the number of observations is very small, similar to the *Unrestricted Communication* treatment, we observe higher trustworthiness for blank messages with delivery than without delivery in the *Restricted Communication* treatment (0 out of 3 without delivery and 3 out 8 with delivery).

Table 2 Unrestricted versus restricted communication—*Roll* rates[ab]

	B's *Roll* rate				
	Unrestricted communication			Restricted communication	Z stat
Condition	Promise	No promise	Total	Total	
Message not delivered	15/43	3/22	18/65	10/22	−1.54
	(35 %)	(14 %)	(28 %)	(45 %)	
Message delivered	25/46	8/19	33/65	11/21	−0.13
	(54 %)	(42 %)	(51 %)	(52 %)	
Z stat	1.84**	2.05**	2.69***	0.45	
Column total	40/89	11/41	51/130	21/43	−1.10
	(45 %)	(27 %)	(39 %)	(49 %)	

[a] The Z stat reflects two sample proportion test for the two populations. [b] In the *Restricted communication* treatment three subjects violated the rules of the pre-play message stage by writing a promise. We exclude the choices made by these subjects from the analysis

*, **, and *** denote significance at $p \leq 0.10$, $p \leq 0.05$, and $p \leq 0.01$ respectively for one tailed test

commitments for those who make a promise, it is just that those who will *Roll* are more likely to make a promise.

Somewhat surprisingly message delivery causes only a small increase in *Roll* rates in the *Restricted communication* treatment. Note that in the *Unrestricted Communication* treatment we found that message delivery increases *Roll* rates even for trustees who do not make a promise. This result seemed to suggest that communication matters irrespective of the content of the messages and one would expect to observe a similar effect of message delivery in the *Restricted Communication* treatment. We are inclined to think that the insignificant effect of message delivery in the *Restricted Communication* treatment is due to the relatively small sample size rather than to a structural behavioral pattern.[7]

[7] This supposition is also supported by the following observation. In the *Restricted communication* treatment we have that 73 % (8/11) of the blank messages were delivered while 41 % (13/32) of the non-blank messages were delivered. As a result the ratio of blank versus non-bank messages is substantially larger in the *Message delivered* condition (8/13 = 0.62) than the in the *Message not delivered* condition (3/19 = 0.16). Since blank messages have lower *Roll* rates than non-blank messages, this tends to bias the *Roll* rates downward among the delivered messages and upward among the undelivered messages. We have simulated how the *Roll* rates in the *Restricted communication* treatment would change if the proportion of delivered messages were close to 50 % for both blank and non-blank messages. Specifically, we assumed that 5 out of the 11 blank messages (45 %) were delivered and 6 out of 11 (55 %) were not. This would mean that 16 of the 32 non-blank messages would be delivered and 16 of the 32 would not be delivered. In this case message delivery would increase *Roll* rates by 21 %, from 36 to 57 %, which is very similar to the observed increase of 23 % in the *Unrestricted communication* treatment (from 28 to 51 %). Moreover, the *Roll* rate for non-delivered messages in the *Restricted communication* treatment (36 %) would be much closer to those observed in the Unrestricted treatment (29 %). Even though, of course, this analysis is somewhat speculative it does support the suggestion that the rather weak effect of communication in the *Restricted communication* treatment may be due to sampling error.

Table 3 Unrestricted versus restricted communication—*In* rates[a]

	A's *In* rate				
	Unrestricted communication			Restricted communication	Z stat
Condition	Promise	No promise	Total	Total	
Message not delivered	20/65		20/65	6/22	0.31
	(31 %)		(31 %)	(27 %)	
Message delivered	34/46	9/19	43/65	7/21	2.65***
	(74 %)	(47 %)	(66 %)	(33 %)	
Z stat	–	–	4.04***	0.43	
Column total	–	–	63/130	13/43	2.09**
			(48 %)	(30 %)	

[a] The Z stat reflects two sample proportion test for the two populations

*, **, and *** denote significance at $p \leq 0.10$, $p \leq 0.05$, and $p \leq 0.01$ respectively for one tailed test

Finally, Table 3 reports results for As. From Table 3 one can see that in the *Unrestricted Communication* treatment As were more likely to play *In* when they received a promise (74 %) than when they received a message with no promise (47 %). In addition, As were more likely to play *In* when they received a message without a promise (47 %) than when they received no message at all (31 %), although this difference is only marginally significant with a one-tailed test (Z stat = 1.34, two sample proportion test, $p < 0.09$, one tailed). Table 3 also compares *In* rates between two treatments by condition. Unlike in the *Unrestricted Communication treatment*, in the *Restricted Communication treatment* A players do not trust B players more when they receive a message than when they do not. Overall, the results reported in Table 3 suggest that A players are more likely to trust in the *Unrestricted communication* treatment than in the *Restricted communication* treatment. Unlike B players, A players are affected by the presence of promises.

4 Discussion

Our results suggest that promises do not cause trustworthiness, they are just more likely to be sent by trustworthy players than by untrustworthy ones. Note though that this does not rule out a preference for internal consistency. Our study does not address why players do (not) make a promise. One possibility is that players do not make a promise due to a reluctance to make a statement that is inconsistent with their prospective behavior. Therefore, as a reviewer suggested, it is possible that the behavior of an untrustworthy trustee would change if a promise could somehow be 'extracted'. Evidence from other experiments, however, speaks against that possibility. Belot et al. (2010) and Charness and Dufwenberg (2010) find that cooperation does not increase with promises that are elicited by a third party. Moreover, in Ismayilov and Potters (2014) we analyze how promise elicitation by A players from B players affects trustworthiness. We find that almost all B players make a promise when asked to do so by the A players. Still, overall trustworthiness is not

higher than in a treatment with one way communication from B player to A player, where promises are volunteered and much less frequent. These results suggest that promises do not increase trustworthiness and do not act as a commitment. Other possible explanations for why some people do not make a promise are that people do not believe that promises will be effective or that making a promise does not come to some (non-cooperative) people's mind. If all people believed that promises are ineffective then the average belief in the *Message delivered* condition of the *Unrestricted Communication* treatment, where promises can be made, would be the same as the average belief in the *Message Delivered* condition of the *Restricted Communication* treatment, where promises cannot be made. In our experiment the average beliefs are higher in the former than in the latter case, but this difference is not significant. (53 vs. 46 %, Z stat = 1.04, p = 0.15 for one-tailed Wilcoxon rank sum test). A caveat of course is that non-cooperative people might believe that making a promise is ineffective and cooperative people might believe it is effective. So even higher average beliefs with unrestricted communication would not necessarily mean that everyone believes that promises are effective in enhancing trust.

It is possible that not making a promise in the *Unrestricted communication* treatment is different from not making a promise in the *Restricted communication* treatment. Then, compared to the *Restricted communication* treatment, allowing for promises in the *Unrestricted communication*, might increase *Roll* rates among those who make a promise (say from 50 to 60 %) and decrease *Roll* rates among those who do not (say from 50 to 40 %). This might be leading to similar aggregate *Roll* rates across treatments, even though behavior is affected by the message sent. Our data cannot rule out this possibility. Still, if promisors feel committed by their message, increasing the rate of promises should increase trustworthiness. However, the evidence discussed in the previous paragraph indicates that an increase in (elicited) promises does not lead to an increase in trustworthiness.

Finally, we find that message delivery matters even for those who do not make a promise in the *Unrestricted Communication* treatment. In fact, even trustees who sent blank messages were more likely to cooperate when their message was delivered than when it was not. These results seem to suggest that in the *Unrestricted Communication* treatment message delivery by itself, irrespective of content, strengthens a feeling of 'closeness' between the trustee and the trustor. The fact that something (a sheet of paper) that was in the trustee's possession is later in the trustor's hands may create some commonality and reduce social distance.[8] However, we do not observe a significant increase in *Roll* rates due to message delivery in the *Restricted Communication* treatment even though the *Roll* rates in the *Restricted Communication* treatment are not lower than in the *Unrestricted Communication* treatment. As mentioned above we are inclined to think that the former result is due to a small sample size in our *Restricted Communication* treatment. The result that the *Roll* rates in the *Restricted Communication* treatment are as high as in the *Unrestricted Communication* treatment is also somewhat surprising given that previous studies have found little evidence for a positive

[8] As one of the referees suggested it may also be that having your message selected for delivery may make you "nicer".

impact of impersonal, game-irrelevant (non-promise) communication on cooperation. For instance, Bouas and Komorita (1996), Mulford et al. (2008), Bicchieri et al. (2010), and He et al. (2014) all find a positive effect of unrestricted communication on cooperation rates relative to no communication or restricted communication, but only He et al. (2014) find that restricted communication (weakly) increases cooperation relative to no communication.[9] We suspect that this might be driven by the fact that communication is two-way in these studies but one-way in our study. Note that in our study we observe significantly higher trust rates by A players with unrestricted communication than with restriction communication when messages are delivered. The existence of promises affects A players, who are the receivers of promises, but not B players, who are the senders of promises. In other words, even though promises by B players do not lead to higher *Roll* rates, they do lead to higher *In* rates. Since with bilateral communication both parties can send and receive promises, players can then be mutually affected by the promise received from the other player (rather than by the promise sent by themselves). This implies that our results are not necessarily incongruent with those of others.

5 Conclusion

This paper set out to test whether promises increase trustworthiness due to a preference for internal consistenty or due to a feeling of social obligation. To address this question we introduced a twist to the trust game of Charness and Dufwenberg (2006) with preplay messages. With 50 % probability a message written by the trustee was not delivered to the trustor. We find that when messages were not delivered trustees who made a promise were more likely to be trustworthy than trustees who did not make a promise. We later ran a control treatment to examine whether this effect was causal in the sense that promises increase trustworthiness. The results do not lend support to this interpretation of promise keeping. It appears that the correlation between promises and trustworthiness is driven by self-selection. We also find that message being delivered increased trustworthiness both by promisors and nonpromisors. This result sheds a new light on the literature, which has largely attributed the effect of communication to the commitment value of promises.

Acknowledgments We thank three anonymous referees and the editor (Jacob Goeree) for helpful comments and suggestions. We also thank Gary Charness and Martin Dufwenberg for filling us in on the details about their procedure, Chen Sun for his help with running the experiments and Marta Serra Garcia, participants at the TIBER seminar at Tilburg University, the 2010 WISE conference at Xiamen University, the M-BEES 2011 at Maastricht University and the 2011 ESA Annual Meeting for helpful comments.

[9] Buchan et al. (2006) show that personal game-irrelevant communication marginally increases trustworthiness relative to impersonal game-irrelevant communication and Roth (1995) finds that personal game-irrelevant face-to-face communication increases average offers and acceptance rates relative to an anonymous no communication treatment in ultimatum games. However, we do not allow subjects to reveal any information that could identify them and anonymity is preserved in our experiment.

References

Belot, M., Bhaskar, V., & van de Ven, J. (2010). Promises and cooperation: Evidence from a TV game show. *Journal of Economic Behavior and Organization*, *73*, 396–405.

Bicchieri, C., Lev-on, A., & Chavez, A. (2010). The medium or the message? Communication relevance and richness in trust games. *Synthese*, *176*, 125–147.

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

Blanco, M., Engelmann, D., Koch, A. K., & Normann, H.-T. (2010). Belief elicitation in experiments: Is there a hedging problem? *Experimental Economics*, *13*, 412–438.

Bouas, K., & Komorita, S. (1996). Group discussion and cooperation in social dilemmas. *Personality and Social Psychology Bulletin*, *22*, 1144–1150.

Buchan, N. R., Johnson, E. J., & Croson, R. T. A. (2006). Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences. *Journal of Economic Behavior and Organisation*, *60*, 373–398.

Charness, G., & Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, *74*, 1579–1691.

Charness, G., & Dufwenberg, M. (2010). Bare promises: An experiment. *Economics Letters*, *107*, 281–283.

Deck, C., Servatka, M., & Tucker, S. (2013). An examination of the effect of messages on cooperation under double-blind and single-blind payoff procedures. *Experimental Economics*, *16*, 597–607.

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

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

He, S., Offerman, T. & van de Ven, J. (2014). Prison talk: Why is communication effective in social dilemmas? Working paper.

Ismayilov, H., & J. Potters (2014). Do people ask for a promise? And should they? Working paper.

Kerr, N., & Kaufman-Gilliland, C. (1994). Communication, commitment, and cooperation in social dilemmas. *Journal of Personality and Social Psychology*, *66*, 513–529.

Lopez-Perez, R., & Spiegelman, E. (2013). Why do people tell the truth? Experimental evidence for pure lie aversion. *Experimental Economics*, *16*, 233–247.

Lundquist, T., Ellingsen, T., & Johannesson, M. (2009). The aversion to lying. *Journal of Economic Behavior and Organization*, *70*, 81–92.

Mulford, M., Jackson, J., & Svedsater, H. (2008). Encouraging cooperation: Revisiting solidarity and commitment effects in prisoners' dilemma games. *Journal of Applied Social Psychology*, *38*, 2964–2989.

Orbell, J. M., van de Kragt, A. J. C., & Dawes, R. M. (1988). Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology*, *54*, 811–819.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *The American Political Science Review*, *86*, 404–417.

Roth, A. (1995). Bargaining experiments. In J. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics*. Princeton, NJ: Princeton University Press.

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society*, *7*, 58–92.

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

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