

ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

*An International Society for the Advancement of Economic
Theory in its Relation to Statistics and Mathematics*

<http://www.econometricsociety.org/>

Econometrica, Vol. 76, No. 6 (November, 2008), 1467–1480

WHY DO PEOPLE KEEP THEIR PROMISES? AN EXPERIMENTAL TEST OF TWO EXPLANATIONS

CHRISTOPH VANBERG

Strategic Interaction Group, Max Planck Institute of Economics, 07745 Jena, Germany

The copyright to this Article is held by the Econometric Society. It may be downloaded, printed and reproduced only for educational or research purposes, including use in course packs. No downloading or copying may be done for any commercial purpose without the explicit permission of the Econometric Society. For such commercial purposes contact the Office of the Econometric Society (contact information may be found at the website <http://www.econometricsociety.org> or in the back cover of *Econometrica*). This statement must be included on all copies of this Article that are made available electronically or in any other format.

WHY DO PEOPLE KEEP THEIR PROMISES? AN EXPERIMENTAL TEST OF TWO EXPLANATIONS¹

BY CHRISTOPH VANBERG

Numerous psychological and economic experiments have shown that the exchange of promises greatly enhances cooperative behavior in experimental games. This paper seeks to test two theories to explain this effect. The first posits that individuals have a preference for keeping their word. The second assumes that people dislike letting down others' payoff expectations. According to the latter account, promises affect behavior only indirectly, because they lead to changes in the payoff expectations attributed to others. I conduct an experiment designed to distinguish between and test these alternative explanations. The results demonstrate that the effects of promises cannot be accounted for by changes in payoff expectations. This suggests that people have a preference for promise keeping per se.

KEYWORDS: Promises, contracts, obligations, guilt aversion, beliefs, behavioral economics, experimental economics.

1. INTRODUCTION

NUMEROUS PSYCHOLOGICAL AND ECONOMIC EXPERIMENTS have shown that pre-play communication, especially the exchange of promises, has profound effects on subsequent levels of cooperation in experimental games (Bicchieri and Lev-On (2007), Charness and Dufwenberg (2006), Ellingsen and Johannesson (2004), Kerr and Kaufmann-Gilliland (1994), Ostrom, Walker, and Gardner (1992), Sally (1995)). This evidence has led many authors to conclude that promises induce emotional commitments to fulfill contractual obligations, perhaps based on a norm of promise keeping (Braver (1995), Ostrom, Walker, and Gardner (1992), Ellingsen and Johannesson (2004)). A formalization of this idea was suggested by Ellingsen and Johannesson (2004), who proposed a model of social preferences that includes a “taste (...) for keeping one's word.”² I will refer to this as the *commitment-based* explanation for promise keeping. Its defining feature is the notion that agents are *directly* concerned, not only about the expected consequences of their behavior, but also about its consistency with *obligations* based on agreements or contracts.

An alternative explanation for promise keeping is based on the theory of *guilt aversion* (Dufwenberg and Gneezy (2000), Battigalli and Dufwenberg (2007)). This theory assumes that human behavior in various social contexts is affected

¹I thank the editor and three anonymous referees for helpful comments and suggestions. I also benefited from discussions with Urs Fischbacher, Werner Gueth, Oliver Kirchkamp, Topi Miettinen, Birendra Rai, Ondrej Rydval, and Anthony Ziegelmeyer. Sascha Brose provided valuable assistance in programming and conducting experimental sessions.

²Similar models have been proposed by, among others, Klein and O'Flaherty (1993) and Miettinen (2008).

by a basic disposition to experience guilt when letting down the payoff expectations attributed to others. In a recent paper, [Charness and Dufwenberg \(2006\)](#) argued that such a disposition may account for promise keeping even if people do not actually care about agreements or contracts per se. According to this explanation, promises are kept because, once given, they cause people to attribute certain payoff expectations to others. I will refer to this as the *expectation-based* explanation for promise keeping.

I will argue that previous experimental evidence does not permit us to distinguish between the commitment-based and expectation-based explanations for promise keeping. To my knowledge, this paper is the first to report on an experiment designed explicitly to distinguish between and test these alternative explanations. The evidence collected demonstrates that the effects of promises cannot be accounted for by changes in payoff expectations. This suggests that people have a preference for keeping promises per se.

The remainder of the paper is organized as follows. Section 2 explains why previous evidence is unable to distinguish between the commitment-based and expectation-based explanations for promise keeping. Section 3 presents the design of the experiment and provides some detail on experimental procedures. Results are reported in Section 4. Section 5 concludes. Appendix A presents additional empirical evidence. Instructions and data are provided online in the Supplemental material ([Vanberg \(2008\)](#)).

2. RELATED LITERATURE

To test the expectation-based account for promise keeping, [Charness and Dufwenberg \(2006\)](#) collected evidence on the effects of promises on both behavior and beliefs in an experimental trust game. In their experiment, the second mover (the “trustee”) is given the opportunity to send a message to the first mover (the “trustor”). Many trustees use this opportunity to make a promise stating that they will act in a trustworthy manner if trusted. The authors then showed that (1) trustees’ second-order beliefs are predictive of trustworthy behavior, both with and without communication, (2) second-order beliefs are shifted by communication, and (3) trustworthiness is higher in the communication treatment. The authors concluded that this evidence supports the notion that communication and promises affect behavior *because* they induce changes in second-order beliefs.

However, as [Charness and Dufwenberg](#) acknowledged, their experiment does not allow them to rule out an alternative explanation for the observed correlation of second-order beliefs and behavior. In particular, *within-treatment* differences in behavior may be driven by heterogeneity in unobserved factors, including altruism or fairness preferences. Differences *between* treatments, in turn, may reflect a preference for promise keeping, per se. In each case,

second-order beliefs may simply *follow* behavior because subjects believe that others can predict their behavior well.³

It follows that [Charness and Dufwenberg's](#) experiment does not allow us to distinguish the commitment-based and expectation-based explanations for promise keeping. The reason is that the causal factor emphasized by the first explanation (the existence of a promise) is perfectly correlated with the one favored by the second explanation (the associated changes in second-order beliefs). The experiment reported on here was designed to achieve *independent* variation in these factors. It therefore allows me to test whether promises themselves have an independent effect on behavior.

3. EXPERIMENTAL DESIGN AND PROCEDURES

3.1. *Experimental Design*

I confront pairs of subjects with a mini dictator game, depicted in Figure 1. At the beginning of the game, Nature randomly determines the identity of the dictator in each pair. This subject then chooses between two actions, *Roll* or *Don't Roll*. If the dictator chooses *Don't Roll*, she receives 14 euros and the recipient receives nothing. If the dictator chooses *Roll*, she receives 10 euros and the recipient receives 12 euros with probability $\frac{5}{6}$ and 0 euros with probability $\frac{1}{6}$. Except for the use of euros rather than dollars, this choice is identical to the one faced by trustees in [Charness and Dufwenberg's \(2006\)](#) trust game.

Prior to Nature's move, subjects are given the opportunity to communicate. Most subjects use this opportunity to exchange promises stating that they will *Roll* if chosen to be the dictator. Consistent with the results obtained by [Charness and Dufwenberg \(2006\)](#), I find that this leads to a significant increase,

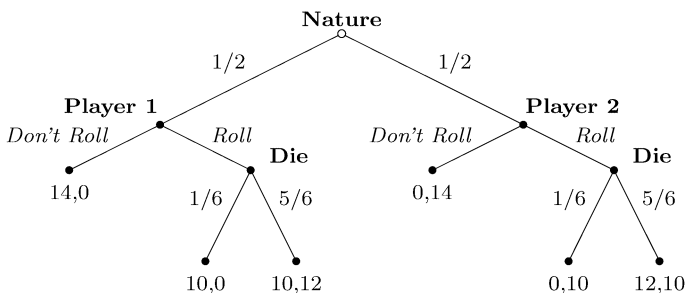


FIGURE 1.—Mini dictator game with random dictatorship.

³This belief has often been referred to as a false consensus effect, a label that is somewhat misleading in that it suggests an irrational bias. As argued by [Dawes \(1989\)](#), however, it can be entirely rational to hold this belief in a world where agents' behavioral inclinations are correlated.

both in the rate at which dictators choose *Roll* and in second-order beliefs concerning the probability that they will do so. These results emerge from a comparison of a no communication and a communication treatment. Results from these treatments are presented in Appendix A.

The goal of the present paper is to investigate whether the increased frequency of *Roll* choices under communication is driven by promise-induced changes in second-order beliefs (as suggested by the expectation-based explanation) or by a preference for promise keeping per se (as suggested by the commitment-based explanation). As indicated above, this question cannot be answered by simply comparing a situation in which subjects communicate to a situation in which they do not. The problem with this approach is that the existence of promises will then be perfectly correlated with any associated changes in beliefs.

The experiment reported on here achieves *independent* variation in promises and second-order beliefs by randomly rematching some subjects after communication takes place. Specifically, recipients in half of all subject pairs are *switched*. While dictator subjects are informed about whether such a switch has occurred, recipients are not.⁴ In case their partners are switched, dictator subjects are given the opportunity to inspect their new recipients' prior conversations. This allows them to learn whether the recipient has received a promise from another participant.

Thus, half of all dictators face a recipient with whom they have communicated. In most cases, this will mean that the dictator has made a promise to the recipient. The other half face *another* recipient. In most cases, this person will have received a promise from someone else. I will refer to the former as the *no switch* condition and the latter as the *switch* condition.

Since recipients do not know whether their partners have been switched, their *first-order beliefs* can depend only on whether a promise was received, not whether they are finally matched with the person who has made that promise. Since dictator subjects know this, their *second-order beliefs* should be equally affected by promises others have made (as in the *switch* condition) as by promises that they themselves have made (as in the *no switch* condition). This hypothesis is verified by the data (see Section 4).

As far as the expectation-based explanation for promise keeping is concerned, the *switch* and *no switch* conditions are therefore identical. Dictator behavior should therefore be equally affected by others' promises as by their own. Thus, the expectation-based explanation for promise keeping predicts that dictators should be more likely to *Roll* whenever the partner has *received* a promise, in both the *switch* and *no switch* conditions.

In contrast, the commitment-based explanation for promise keeping suggests that a promise establishes a contractual obligation toward the person to

⁴Naturally, all subjects know that the probability of being switched is $\frac{1}{2}$. Thus, the experiment does not involve deception.

whom that promise is made (i.e., the chatting partner). This implies that a dictator's behavior should be affected only by her *own* promise and only in the *no switch* condition.⁵ Thus, the commitment-based explanation predicts that dictators should be more likely to choose *Roll* in the *no switch* condition.⁶

3.2. *Experimental Procedure*

The experiment was conducted in the experimental laboratory of the Max Planck Institute of Economics in Jena, Germany. Subjects were students from the Friedrich-Schiller University of Jena. Subjects were recruited using the online recruitment system ORSEE (Greiner (2004)). Upon entering the laboratory, subjects were randomly assigned to 32 visually isolated computer terminals. Instructions (reproduced in the Supplemental material as Appendix C) were distributed and read out loud by the experimenter (myself). Questions were answered individually at the subjects' seats. Before beginning the experiment, subjects filled out a short questionnaire testing comprehension of the rules.

Each session consisted of 8 rounds, with perfect stranger matching.⁷ After the experiment, one of the 8 rounds was chosen for payment of the decision. Elicitation of beliefs was incentivized with modest payoffs. The precise procedure is described in more detail below. Beliefs were paid in all rounds except the one chosen for payment of the decision.⁸ All subjects received a fixed participation fee of 2.50 EUR. The experiment was programmed and conducted with the software z-Tree (Fischbacher (2007)).

The sequence of events in each round was as follows. First, the 32 subjects were randomly matched to form 16 *chatting pairs*. Beginning with a randomly chosen participant, subjects alternated sending messages. Each person could send two messages of at most 90 characters. During the communication phase, neither subject knew which member of the pair would subsequently be a dictator.

After communication, one subject in each chatting pair was randomly chosen to be a dictator. Next, half of all chatting pairs (i.e., 8) were randomly assigned

⁵The assumption that a promise establishes a contractual obligation only toward the chatting partner (rather than a more general obligation) biases the analysis in favor of the expectation-based explanation for promise keeping.

⁶This design should minimize a correlation of second-order beliefs with the existence of a promise. However, the "false consensus" argument mentioned above implies that we can still expect a correlation of second-order beliefs and behavior *within* each of the conditions. Such a correlation may reflect the influence of guilt aversion or of other unobserved factors.

⁷This matching scheme guarantees that no two subjects meet more than once (either to communicate *or* to interact). This implies that each round is strategically equivalent to a one shot situation.

⁸This procedure implies that subjects have no incentive to misreport beliefs in an attempt to hedge against bad outcomes.

Subject A Will...	Certainly Choose <i>Roll</i>	Probably Choose <i>Roll</i>	Unsure	Probably Choose <i>Don't Roll</i>	Certainly Choose <i>Don't Roll</i>
Your guess	0	0	0	0	0
Your earnings if the other participant chooses <i>Roll</i>	0.65 EUR	0.60 EUR	0.50 EUR	0.35 EUR	0.15 EUR
Your earnings if the other participant chooses <i>Don't Roll</i>	0.15 EUR	0.35 EUR	0.50 EUR	0.60 EUR	0.65 EUR

FIGURE 2.—Elicitation of first-order beliefs.

to the *switch* and half to the *no switch* condition. Dictators in the *no switch* condition remained with their chatting partners during the subsequent decision phase. Dictators in the *switch* condition were matched with a new recipient from another pair.

Prior to the decision phase, subjects saw a screen informing them of their role. In addition, dictator subjects were told whether or not their partner had been switched. If so, dictators could click a button to view the content of their new partners' previous conversations.⁹ Recipients were informed only of their role, not whether their partners were switched.

Next, dictator subjects were asked to submit their decisions. At the same time, recipients were asked to guess the decision of the unknown person to whom they were now matched. Specifically, recipients were asked to choose a column from the five point scale depicted in Figure 2. Each column is associated with payoffs that depend on the decision made by the partner. This procedure yields a five point scale for *first-order beliefs*.¹⁰

After decisions and first-order beliefs were elicited, dictator subjects were asked to guess the recipient's guess concerning their own behavior. Specifically, dictators were presented with the table depicted in Figure 2 and asked to mark the box that they believed the recipient had clicked. If they guessed correctly, dictators earned 0.50 EUR. This yields a five point scale for *second-order beliefs*.

At the end of each round, subjects received feedback about the payoffs they would receive if the current round was chosen for payment. The roll of the die

⁹This guaranteed that a dictator could find out whether the recipient had received a promise. The button was clicked in 96% of all cases. All subjects (including recipients) could click a button to review their own chat. This guaranteed that switched dictators could not be identified by mouse clicks. To ensure that dictators could differentiate between chatting partners and interaction partners, all subjects were assigned neutral identification numbers (e.g., R9). These were used when displaying chat contents and payoff tables.

¹⁰The payoffs correspond to a quadratic scoring rule for the probabilities 85%, 68%, 50%, 32%, and 15%. Because it is based on the assumption of risk neutrality, the quadratic scoring yields notoriously flat payoffs as probabilities approach certainty. For this reason, I limited the available choices to less extreme probabilities. For illustration, beliefs are later scaled to 0, $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$, and 1.

was simulated using z-Tree's random number generator. Dictator subjects who had chosen to *Roll* were told the number that had been rolled. Recipients were told only their payoff. Recipients were not informed of the dictator's choice or whether their partner had been switched. Feedback regarding the payoffs from guessing was given only at the very end of the experiment.

4. RESULTS

The data comprise 6 experimental sessions involving a total of 192 subjects. Each session lasted for 8 rounds. The average number of dictator decisions made by each subject is 4. Half of these decisions were made in the *switch* condition and half in the *no switch* condition. Thus, we have a total of 384 decisions made under each condition. However, each session of the experiment constitutes only one independent observation. Nonparametric tests are therefore based on session averages of the relevant variables. Panel regressions making use of the individual-level data are reported in Appendix B.

4.1. *The Effects of Partner Switching*

Table I summarizes dictators' second-order beliefs and behavior under the *switch* and *no switch* conditions. The *Z* statistics and *p* values reflect Wilcoxon signed-rank tests using session-level averages of second-order beliefs and *Roll* rates. As expected, dictators' second-order beliefs did not differ significantly between the *switch* and *no switch* conditions. Nonetheless, dictators were significantly more likely to *Roll* when interacting with a person with whom they themselves had communicated than when interacting with someone who had communicated with another subject.

This result is inconsistent with the idea that the effects of communication are driven by changes in second-order beliefs. If subjects exchanged promises to *Roll* during the communication phase, a possible interpretation is that dictators were more inclined to keep their *own* promises (as in the *no switch* condition), but not those that *others* had made (as in the *switch* condition), even though

TABLE I
THE EFFECTS OF PARTNER SWITCHING^a

	No Switch	Switch	Z Stat
Average second-order belief	0.76 (0.26)	0.71 (0.28)	<i>Z</i> = 1.48 (<i>p</i> = 0.14)
Fraction of subjects choosing <i>Roll</i>	69% (0.02)	54% (0.03)	<i>Z</i> = 2.20 (<i>p</i> = 0.03)

^aPooled data from all sessions, all rounds. Each cell *N* = 384. Standard errors in parentheses. The *Z* statistic reflects Wilcoxon signed-rank tests using session level *Roll* rates and average second-order beliefs.

they knew that recipients were unable to distinguish between the two when forming their expectations.

4.2. *The Effects of Promises*

To investigate this interpretation more directly, I asked a student assistant to code messages according to whether or not they contained a promise or statement of intent stating that the subject would *Roll*.¹¹ Overall, 79% of messages sent contained such a promise. Using this classification, Tables II and III summarize dictator beliefs and behavior under 6 different conditions. The top row of each table reports average beliefs (resp. roll rates) among dictators who promised to *Roll* during the communication phase. The bottom row reports values for those who *did not* make such a promise. The columns distinguish between dictators whose partners were *not switched* (column 1) and those whose partners were *switched* (columns 2 and 3). In addition, columns 2 and 3 differentiate between (new) recipients who did (column 2) or did not (column 3) receive a promise from another subject.

Within the *no switch* condition, dictators who promised to *Roll* did so in 73% of all cases. This fraction is significantly larger than the 52% *Roll* rate among those who did not make such promises ($Z = 2.2$, $p = 0.03$). Similarly, dictators who promised and were not switched also had significantly higher second-order beliefs than those who did not promise and were not switched ($Z = 2.2$, $p = 0.03$). This pattern mirrors the results of Charness and Dufwenberg (2006), and it is consistent with both the expectation-based and commitment-based explanations for promise keeping.

However, the fraction of dictators who chose *Roll* after promising and not switching partners is also significantly larger than the 54% *Roll* rate among

TABLE II
PROMISES AND SECOND-ORDER BELIEFS^a

	Average Second-Order Belief (All Sessions)		
	No Switch	Switch	
		Partner Received a Promise	Partner Received No Promise
Dictator promised	0.80 (0.23)	0.76 (0.25)	0.62 (0.31)
Dictator did not promise	0.60 (0.33)	0.70 (0.31)	0.58 (0.29)

^a Pooled data from all sessions, all rounds. Standard errors in parentheses.

¹¹ More specifically, each *pair* of messages sent by a particular subject in a particular round was treated as one message.

TABLE III
PROMISES AND BEHAVIOR^a

	Fraction of Subjects Choosing Roll (All Sessions)		
	No Switch	Switch	
		Partner Received a Promise	Partner Received No Promise
Dictator promised	227/309 (73%) [0.03]	129/238 (54%) [0.03]	29/56 (52%) [0.07]
Dictator did not promise	39/75 (52%) [0.06]	30/56 (54%) [0.07]	19/34 (56%) [0.09]

^aPooled data from all sessions, all rounds. Standard errors in brackets.

those who promised and were then rematched to interact with a partner who had received a promise from another subject ($Z = 2.2$, $p = 0.03$). This difference in behavior occurred despite the fact that second-order beliefs did not differ significantly between these conditions ($Z = 1.57$, $p = 0.12$).

These results indicate that a dictator's promise affects behavior toward the specific person to whom that promise is made.¹² Contrary to the expectation-based explanation for promise keeping, this effect does not appear to be driven by a general motivation to fulfill the expectations that a promise induces in the person who has received it.

Further evidence in support of this conclusion is provided by the *switch* condition. Here, dictators whose partners had received promises from someone else had significantly higher second-order beliefs than those whose partners had not received such promises.¹³ The expectation-based explanation for promise keeping would therefore predict that promises received should also lead to significantly higher *Roll* rates. However, this prediction is clearly falsified by the data. Irrespective of whether or not the dictator had made a promise, promises received by the partner were not associated with significant differences in the likelihood of choosing *Roll*.¹⁴

¹²A dictator's own promise is not predictive of behavior toward others in the *switch* condition ({promised, promise received} vs. {didn't promise, promise received}: $Z = 0.94$, $p = 0.35$; {promised, no promise received} vs. {didn't promise, no promise received}: $Z = 0.42$, $p = 0.67$).

¹³{promised, promise received} vs. {promised, no promise received}: $Z = 2.2$, $p = 0.03$; {didn't promise, promise received} vs. {didn't promise, no promise received}: $Z = 1.94$, $p = 0.05$.

¹⁴{promised, promise received} vs. {promised, no promise received}: $Z = 0.52$, $p = 0.60$; {didn't promise, promise received} vs. {didn't promise, no promise received}: $Z = 0.31$, $p = 0.75$).

4.3. *Summary of Results*

To summarize, second-order beliefs did not differ significantly between the *switch* and *no switch* conditions. Nonetheless, dictators were significantly more likely to *Roll* in the *no switch* condition. While a dictator's *own* promise significantly affects both beliefs and behavior toward the person to whom a promise was made, promises made by others lead to significant changes in second-order beliefs, but do not affect behavior.

These results appear to be incompatible with the expectation-based explanation for promise keeping, according to which promise keeping behavior is driven by changes in second-order beliefs. The fact that a dictator's *own* promise has an effect on behavior toward a particular person lends support to the commitment-based explanation for promise keeping, according to which a promise creates a contractual *obligation* toward the person to whom it is made.

5. CONCLUSION

Numerous psychological and economic experiments have shown that the exchange of promises greatly enhances cooperative behavior in experimental games. The aim of this paper was to distinguish between and test two explanations for this effect. The commitment-based explanation says that promises affect behavior *directly*, because people have a preference for keeping their word. The expectation-based explanation says that promises affect behavior *indirectly*, because they lead to changes in second-order beliefs, which in turn affect behavior.

Previous experiments have not been able to distinguish between these explanations, because treatment-induced changes in the alternative causal factors (promises and second-order beliefs) have occurred simultaneously. The experiment reported on here was designed to achieve independent variation in promises and promise-induced changes in second-order beliefs. The evidence demonstrates that a dictator's *own* promise significantly affects behavior toward the person to whom a promise was made. In contrast, promises made by others lead to significant changes in second-order beliefs, but do not affect behavior. This evidence shows that promises, *per se*, have an independent effect on behavior, as suggested by the commitment-based account.

Before concluding, it may be useful to comment briefly on the concept of *guilt* and its role in the theories under discussion. Although the expectation-based account was explicitly based on the theory of *guilt aversion*, it should be emphasized that the commitment-based explanation is entirely compatible with the notion that promise keeping is ultimately motivated by a desire to avoid a negative emotion, which we might as well refer to as guilt. What is at issue is not whether the behavior we are seeking to explain is driven by such a desire. Instead, the relevant question is what conditions *trigger* feelings of guilt? It is in their answer to this question that the two theories discussed above differ.

The theory of guilt aversion assumes that guilt is triggered by behavior that leads to an outcome which is inconsistent with others' empirical *expectations* concerning *what will happen*. The expectation-based explanation for promise keeping holds that this is why people feel guilty when they break a promise.¹⁵ In contrast, the commitment-based explanation holds that people feel guilty about breaking promises, per se. Here, guilt is triggered by behavior that is inconsistent with contractual or moral *obligations* concerning *what "should" be done*.

On a more general level, the evidence reported on in this paper might be interpreted as supporting a kind of "behavioral contractarianism," according to which a subject's appraisal of a situation can be directly affected by her perception of the rights and obligations that apply. This concept constitutes a significant deviation from the theory of guilt aversion as currently proposed. Future attempts to account for the influence of contracts and norms should acknowledge and explain the independent effects that these sources of obligation can have on human motivation.

APPENDIX A: BASELINE TREATMENTS

I ran two baseline treatments to test whether the results of [Charness and Dufwenberg \(2006\)](#) can be replicated using the mini dictator game depicted in Figure 1. In a no communication treatment, subjects were not allowed to communicate at all. In a communication treatment, subjects communicated prior to interacting, and partners remained fixed thereafter. To obtain a larger number of observations from a single session, the strategy method was used.¹⁶

Table A.I presents summary data on behavior and beliefs in the communication and no communication treatments, as well as *Z* statistics reflecting a test of proportions (for roll rates) and a Wilcoxon rank sum test (for beliefs) using the 32 independent first round observations under each treatment. Consistent with [Charness and Dufwenberg \(2006\)](#), I find that communication significantly affected *both* the roll rate and second-order beliefs.

¹⁵A similar argument can be made concerning norm-abiding behavior more generally. [Charness and Dufwenberg \(2006, p. 1596\)](#) noted that "one central idea [in the literature on social norms] is to view a social norm as a moral expectation, which people are inclined to live up to. We suggest that in many cases guilt aversion can provide a kind of microfoundation for this. (...) There is a norm, it shapes [A's] expectation, and [B] lives up to this expectation because he would feel guilty if he did not."

¹⁶In the main treatment with partner switching, using the strategy method would have required me to explicitly ask subjects to submit a strategy that conditions on whether their partner had been switched, possibly introducing a significant demand effect. Previous research has shown that decisions elicited using the strategy method do not differ significantly from those elicited using this "hot" method ([Brandts and Charness \(2000\)](#)). Nonetheless, I will not make direct comparisons between the baseline treatments and the main treatment.

TABLE A.I
THE EFFECTS OF COMMUNICATION^a

	No Communication	Communication	Z Stat
Mean second order belief	0.45 (0.27)	0.92 (0.24)	Z = 6.07 ($p < 0.01$)
Fraction of subjects choosing <i>Roll</i>	51% (0.03)	74% (0.03)	Z = 2.46 ($p = 0.01$)

^aPooled data from all rounds. Each cell $N = 256$. Standard errors in parentheses. The Z statistic reflects a test of proportions (for roll rates) and a Wilcoxon rank sum test (for beliefs), using first round observations only.

APPENDIX B: PANEL REGRESSIONS

The nonparametric tests reported in the main text were based on session level averages. In this section, I present additional results from estimating panel models that make use of individual-level data. Table A.II presents estimates of random intercept binomial logit models. Beginning from a very general model (column 1), I eliminate insignificant interaction effects, finally arriving at a relatively simple model, presented in column 4 (RILOG4).¹⁷

These estimation results are in line with the conclusions derived in the main text. Specifically, dictator subjects who issued promises were significantly more likely to *Roll*. However, this effect was observed only if their partners were not switched. Promises *received* by the partner did not have a significant impact on behavior when partners were switched.

Note that the coefficient on second-order beliefs is highly significant in all regressions. As indicated in the [Introduction](#), this result may reflect a direct *causal* impact of second-order beliefs on behavior, as suggested by the theory of guilt aversion. However, it may also reflect a correlation of second-order beliefs with unobserved factors, based on a false consensus effect. For the purposes of the question addressed in this paper, what is important is that the effects of promises and partner switching are significant while second-order beliefs are controlled for. This indicates that promises *themselves* affect behavior in a way that cannot be accounted for by the expectation-based theory of promise keeping.

REFERENCES

- BATTIGALLI, P., AND M. DUFWENBERG (2007): "Guilt in Games," *American Economic Review, Papers and Proceedings*, 97, 170–176. [1467]
- BICCHIERI, C., AND A. LEV-ON (2007): "Computer-Mediated Communication and Cooperation in Social Dilemmas: An Experimental Analysis," *Politics, Philosophy, and Economics*, 6, 139–168. [1467]

¹⁷Estimates from the logit model are similar to those obtained using alternative methods such as a probit, a conditional logit, or a random effects linear probability model.

TABLE A.II
ESTIMATES OF PANEL REGRESSIONS^a

	RILOG1	RILOG2	RILOG3	RILOG4	PROBIT	CLOGIT	LINEAR
Dictator promised to <i>Roll</i>	2.87*** (0.99)	3.33*** (1.18)	3.37** (1.32)	1.39*** (0.39)	0.78*** (0.22)	1.36** (0.54)	0.16*** (0.05)
Partner was switched	0.12 (1.00)	0.41 (1.00)	0.24 (0.58)	0.44 (0.52)	0.23 (0.30)	0.63 (0.55)	0.05 (0.07)
Second-order belief	5.39*** (1.49)	5.90*** (1.65)	5.63*** (1.16)	3.52*** (0.45)	1.96*** (0.24)	2.68*** (0.62)	0.43*** (0.05)
Dictator promised × partner switched	-1.09 (0.72)	-1.72** (0.80)	-1.71** (0.81)	-1.86** (0.76)	-1.04** (0.45)	-2.03*** (0.64)	-0.22** (0.10)
Dictator promised × second-order belief	-2.31 (1.69)	-3.01* (1.79)	-3.00 (1.88)				
Partner switched × partner received promise	-0.43 (1.54)	0.35 (0.42)					
Partner switched × second-order belief	0.24 (1.33)	-0.60 (1.21)					
Dictator promised × partner switched × partner received promise	0.69 (1.28)						
Partner switched × partner received promise × second-order belief	0.47 (1.51)						
Dictator promised × partner switched × second-order belief	-1.57 (1.92)						
Round	-0.21*** (0.08)	-0.21*** (0.08)	-0.21*** (0.07)	-0.19*** (0.07)	-0.11*** (0.04)	-0.17*** (0.05)	-0.02** (0.01)
Constant	-2.37** (1.04)	-2.69** (1.17)	-2.57*** (0.78)	-1.36*** (0.51)	-0.76*** (0.28)		0.33*** (0.06)
Log likelihood	-385.18	-385.70	-386.19	-389.79	-390.76	-111.33	

^aThe dependent variable is the probability that the dictator chooses *Roll*. Columns 1–5 are random intercept logit and probit models. These models are estimated using GLLAMM (Stata). Standard errors take into account clustering by session. Column 6 is a conditional logit and column 7 is a random effects linear probability model. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

- BRANDTS, J., AND G. CHARNES (2000): "Hot vs. Cold: Sequential Responses and Preference Stability in Experimental Games," *Experimental Economics*, 2, 227–238. [1477]
- BRAVER, S. (1995): "Social Contracts and the Provision of Public Goods," in *Social Dilemmas: Perspectives on Individuals and Groups*, ed. by D. Schroeder. New York: Praeger, 69–86. [1467]
- CHARNES, G., AND M. DUFWENBERG (2006): "Promises and Partnership," *Econometrica*, 74, 1579–1601. [1467-1469, 1474, 1477]
- DAWES, R. (1989): "Statistical Criteria for Establishing a Truly False Consensus Effect," *Journal of Experimental Social Psychology*, 25, 1–17. [1469]
- DUFWENBERG, M., AND U. GNEEZY (2000): "Measuring Beliefs in an Experimental Lost Wallet Game," *Games and Economic Behavior*, 30, 163–182. [1467]
- ELLINGSEN, T., AND M. JOHANNESSON (2004): "Promises, Threats and Fairness," *The Economic Journal*, 114, 397–420. [1467]

- FISCHBACHER, U. (2007): "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments," *Experimental Economics*, 10, 171–178. [1471]
- GREINER, B. (2004): "An Online Recruitment System for Economic Experiments," in *Forschung und Wissenschaftliches Rechnen 2003*, ed. by K. Kremer and V. Macho. GWDG Bericht 63. Gesellschaft für Wissenschaftliche Datenverarbeitung Göttingen: Datenverarbeitung, 79–93. [1471]
- KERR, N., AND C. KAUFMANN-GILLILAND (1994): "Communication, Commitment, and Cooperation in Social Dilemmas," *Journal of Personality and Social Psychology*, 66, 513–529. [1467]
- KLEIN, D., AND B. O'FLAHERTY (1993): "A Game-Theoretic Rendering of Promises and Threats," *Journal of Economic Behavior and Organization*, 21, 295–314. [1467]
- MIETTINEN, T. (2008): "Contracts, Promises and Norms—An Approach to Pre-Play Agreements," Working Paper, Max Planck Institute of Economics. [1467]
- OSTROM, E., J. WALKER, AND R. GARDNER (1992): "Covenants With and Without the Sword: Self-Governance Is Possible," *American Political Science Review*, 86, 404–417. [1467]
- SALLY, D. (1995): "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments From 1958 to 1992," *Rationality and Society*, 7, 58–92. [1467]
- VANBERG, C. (2008): "Supplement to 'Why Do People Keep Their Promises? An Experimental Test of Two Explanations'," *Econometrica Supplementary Material*, 76, http://www.econometricsociety.org/ecta/Supmat/7673_instructions_to_experimental_subjects-1.pdf; http://www.econometricsociety.org/ecta/Supmat/7673_instructions_to_experimental_subjects-2.pdf; http://www.econometricsociety.org/ecta/Supmat/7673_data_and_programs.zip. [1468]

Strategic Interaction Group, Max Planck Institute of Economics, Kahlaische Str. 10, 07745 Jena, Germany; vanberg@econ.mpg.de.

Manuscript received January, 2008; final revision received May, 2008.