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

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, and especially the exchange of promises, have profound effects on subsequent levels of cooperation in experimental games (Bicchieri and Lev-On, Charness and Dufwenberg, Ellingsen and Johannesson, Kerr and Kaufmann-Gilliland, Ostrom, Walker, and Gardner, Sally). 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, Ostrom, Walker, and Gardner, Ellingsen and Johannesson). A formalization of this idea has been suggested by Ellingsen and Johannesson, who propose a model of social preferences that includes a “taste (...) for keeping one’s word.” 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, Battigalli and Dufwenberg). This theory assumes that human behavior in various social contexts is affected by a basic disposition to experience guilt when letting down the payoff expectations attributed to others. In a recent paper, Charness and Dufwenberg

---

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).
argue that such a disposition may account for promise keeping even if people do not actually care about agreements or contracts, \emph{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 \emph{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, \emph{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

In order to test the expectation-based account for promise keeping, Charness and Dufwenberg (2006) collect 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 \emph{promise} stating that they will act in a trustworthy manner if trusted. The authors then show 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 conclude that this evidence supports the notion that communication and promises affect behavior \emph{because} they induce changes in second-order beliefs.

However, as Charness and Dufwenberg acknowledge, their experiment does not allow them to rule out an alternative explanation for the observed correlation of second-order beliefs and behavior. In particular, \emph{within-treatment} differences in behavior may be driven by heterogeneity in unobserved factors, including altruism or fairness preferences. Differences \emph{between} treatments, in turn, may reflect a preference for promise keeping, \emph{per se}. In each case, second-order beliefs may simply \emph{follow} behavior because subjects believe that others can predict their behavior well.\footnote{This 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...}
WHY DO PEOPLE KEEP THEIR PROMISES?

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, 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.

correlated.
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. 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 are 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 towards the person to 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. Thus, the commitment-based explanation predicts that dictators should

---

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

4 The assumption that a promise establishes a contractual obligation only towards the chatting partner (rather than a more general obligation) biases the analysis in favor of the expectation-based explanation for promise keeping.
be more likely to choose *Roll* in the *no switch* condition.\(^5\)

### 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 of 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 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.\(^6\) 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.\(^7\) 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 in 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 will 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 to the *switch*, and half to the *no switch* condition. Dictators in the *no switch* condition remained with their chatting partner 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 partner’s previous conversations.\(^8\) Recipients were informed only of their role.

---

\(^5\)This 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.

\(^6\)This 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.

\(^7\)This procedure implies that subjects have no incentive to misreport beliefs in an attempt to hedge against bad outcomes.

\(^8\)This guaranteed that a dictator could find out whether the recipient had received a promise.
not whether their partner was 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 are now matched. Specifically, recipients were asked to choose a column from the five point scale depicted in Figure 2. Each column is associated with a prospect whose payoff depends on the decision made by the partner. This procedure yields a five point scale for first-order beliefs.9

<table>
<thead>
<tr>
<th>Subject A will...</th>
<th>certainly Roll</th>
<th>probably Roll</th>
<th>unsure</th>
<th>probably Not Roll</th>
<th>certainly Not Roll</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your guess:</td>
<td>0</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>Your earnings if the other participant chooses Roll</td>
<td>0.65 EUR</td>
<td>0.60 EUR</td>
<td>0.50 EUR</td>
<td>0.35 EUR</td>
<td>0.15 EUR</td>
</tr>
<tr>
<td>Your earnings if the other participant chooses Don’t Roll</td>
<td>0.15 EUR</td>
<td>0.35 EUR</td>
<td>0.50 EUR</td>
<td>0.60 EUR</td>
<td>0.65 EUR</td>
</tr>
</tbody>
</table>

Figure 2.— Elicitation of first-order beliefs

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 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, nor 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

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. In order 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.

9The 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.
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. None the less, 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 they knew that recipients were unable to distinguish between the two when forming their expectations.

4.2. The effects of promises

In order to investigate this interpretation more directly, I asked a student assistant to code messages according to whether or not they contain a promise or statement of intent stating that the subject will Roll.\textsuperscript{10} 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

\textsuperscript{10}More specifically, each pair of messages sent by a particular subject in a particular round was treated as one message.
Table II

Promises and second-order beliefs

<table>
<thead>
<tr>
<th></th>
<th>No switch</th>
<th>Partner received a promise</th>
<th>Partner received no promise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dictator promised</td>
<td>0.80</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.31)</td>
<td></td>
</tr>
<tr>
<td>Dictator did not promise</td>
<td>0.60</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(0.29)</td>
<td></td>
</tr>
</tbody>
</table>

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

Table III

Promises and behavior

<table>
<thead>
<tr>
<th></th>
<th>No switch</th>
<th>Partner received a promise</th>
<th>Partner received no promise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dictator promised</td>
<td>227/309</td>
<td>129/238</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(73%)</td>
<td>(54%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.03]</td>
<td>[0.07]</td>
<td></td>
</tr>
<tr>
<td>Dictator did not promise</td>
<td>39/75</td>
<td>19/34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(52%)</td>
<td>(56%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.06]</td>
<td>[0.09]</td>
<td></td>
</tr>
</tbody>
</table>

Pooled data from all sessions, all rounds. Standard errors in brackets.

The 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 explanation for promise keeping.

However, the fraction of dictators choosing Roll after promising and not switching partners is also significantly larger than the 54% Roll rate among 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 \((Z = 1.57, p = 0.12)\).
WHY DO PEOPLE KEEP THEIR PROMISES?

These results indicate that a dictator’s promise affects behavior towards the specific person to whom that promise is made.\textsuperscript{11} 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.\textsuperscript{12} 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.\textsuperscript{13}

4.3. Summary of results

To summarize, second-order beliefs did not differ significantly between the switch and no switch conditions. None the less, dictators were significantly more likely to Roll in the no switch condition. While a dictator’s own promise significantly affects both beliefs and behavior towards 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 towards a particular person lends support to the commitment-based explanation for promise keeping, according to which a promise creates a contractual obligation towards 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 be-

\textsuperscript{11} A dictator’s own promise is not predictive of behavior towards 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\}: \(Z = 0.42, p = 0.67\)).

\textsuperscript{12} \{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\}: \(Z = 1.94, p = 0.05\).

\textsuperscript{13} \{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\}: \(Z = 0.31, p = 0.75\).
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 second-order beliefs. The evidence demonstrates that a dictator’s own promise significantly affects behavior towards 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. 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 subjects’ 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.

Max Planck Institute of Economics, Kahlaische Str. 10, 07745 Jena, Germany

14 A similar argument can be made concerning norm-abiding behavior more generally. Charness and Dufwenberg (2006, p.1596) note 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.”
WHY DO PEOPLE KEEP THEIR PROMISES?

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. In order to obtain a larger number of observations from a single session, the strategy method was used.\textsuperscript{15}

Table A1 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.

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 A2 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).\textsuperscript{16}

\textsuperscript{15}In 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)). None the less, I will not make direct comparisons between the baseline treatments and the main treatment.

\textsuperscript{16}Estimates from the logit model are similar to those obtained using alternative methods, such as a probit, conditional logit, or random effects linear probability model.
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.
WHY DO PEOPLE KEEP THEIR PROMISES?

Table A.2: Estimates of Panel Regressions

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

***significance at 1%; **significance at 5%; *significance at 10%. The dependent variable is the probability that the dictator chooses Roll. Columns 1 to 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 RE linear probability model.
REFERENCES


