PRINCIPIA LXV (2018): 5–33
PL-ISSN 0867-5392
DOI 10.4467/20843887PI.18.010.10869

Florian Ederer, Alexander Stremitzer

Moral Intuitions of Promise Keeping

Intuicje moralne na temat dotrzymywania obietnic

Summary

Promises are a pervasive and important feature of real-world economic exchange situations. We investigate lay people's intuitions of promise keeping. We study the effect of mutual promises, the dynamic of promising and performance over time, the effect of continuous as opposed to binary performance decisions, the effect of income, and the role the receipt of a promise plays on promise-keeping. Assuming that law serves as a backstop of moral intuitions, our results cast some light on the mutuality requirement, the doctrine of substantial performance, doctrines of divisible obligations, and doctrines of contract formation.

Keywords: contracts, promises, moral intuitions, mutuality, law
JEL-Classification: K12, A13, C91, C72

Streszczenie

Składanie obietnic jest powszechnym i istotnym zjawiskiem w realnych sytuacjach obrotu gospodarczego. W niniejszym artykule badamy intuicje laików na temat dotrzymywania obietnic. Śledzimy efekt wzajemnych przyrzeczeń, dynamikę składania i spełniania obietnic na przestrzeni czasu. Badamy zagadnienie rozciągniętego w czasie spełniania obietnic w przeciwieństwie do pojedynczych, wzajemnych decyzji; badamy wpływ wysokości dochodu i to, jaką rolę odgrywa fakt uzyskania obietnicy w jej dotrzymaniu. Zakładając, że prawo jest narzędziem wspierania intuicji moralnych, rezultaty naszych badań rzucają światło na wymóg wzajemności, na teorię zasadniczego spełnienia, podzielności zobowiązań i na teorię zawierania umów.

Słowa kluczowe: umowy, obietnice, intuicje moralne, wzajemność, prawo
Klasyfikacja JEL: K12, A13, C91, C72

0. Introduction

Promises are a pervasive and important feature of real-world economic exchange situations. Until recently, however, the mechanisms behind promise-keeping have remained relatively neglected by academic research in economics, psychology, philosophy, and law. Nevertheless, a clear account of what drives commitment to promises is essential to harness those effects in institutional design, whether it be in the design of legal policy, of regulatory regimes, or of contracts and organizations.

If legal and institutional regimes track and reinforce moral intuitions, we would assume that by studying the mechanisms underlying promise-keeping, we could also learn much about the structure of legal rules. Economists think of promises as unbinding preplay communication, as opposed to enforceable contracts, and have shown that such communication influences performance decisions. In other words, promises allow parties to engage in mutually beneficial interactions even in the absence of extrinsic sanctions (Charness and Dufwenberg 2006).

Contracts in a legal sense are "legally enforceable promises." Not all promises are contracts, but all contracts are promises. Therefore, two sources might motivate people to keep their promises: the sanctions for breach of contract afforded by an external enforcement mechanism, like formal or reputational sanctions, and an intrinsic sense of obligation fostered by the moral force of being bound by a promise. Traditionally, debates on promises and reasons to keep them were the purview of philosophers, while discussing external enforcement was the purview of economists. However, the advent of behavioral and experimental economics and the more recent advent of experimental philosophy have partially bridged this divide and have also blurred the borders with neighboring disciplines such as psychology and neuroscience.

Economists have studied in great detail how to structure externally enforced incentive schemes to achieve desired outcomes. Perhaps more importantly, economists have derived impossibility results which show that in a setting of bilateral asymmetric information, it is impossible to achieve socially optimal exchange (see, e.g., Myerson and Satterthwaite 1983; Che and Hausch 1999). However, if it is possible to enlist the moral force of promises using institutional arrangements, these impasses may be overcome. On the other hand, philosophers have developed sophisticated arguments detailing good reasons for keeping promises (for an overview see, e.g., Shiffrin 2008). However, a detailed and systematic analysis of the mechanisms driving the moral force of promise-keeping is lacking. Such a theory would be needed to inform institutional design, which would harness this moral force.

In this paper we analyze the effect of mutuality in promises. Moreover, we investigate the dynamics of promising, cooperation, and promise-keeping in order to see whether legal rules, by encouraging the exchange of commitments, contribute to the breeding of cooperation over time. We also study how the introduction of continuous breach, as opposed to binary breach, affects levels and rates of performance. Another topic of interest is how promise-keeping depends on income. Finally, we study whether people are more likely to act on a promise if the promise was received by the promisee. There could be two reasons for that to be the case: Either the fact that the message is not received weakens the promise *per se*, or it reduces the guilt from promise-breaking, as a promisee who did not receive the promise did not expect performance. We try to make a step toward disentangling those two motivations.

On the effect of mutuality, we find that promisors do not reward explicit counter-promises. However, they punish counterparties who explicitly refuse to make a counter-promise. We also find that the possibility of mutual promising increases promising over time but does not increase promise-keeping. Promise-keeping decreases with income, although other-regarding preferences do not depend on income. We see that the likelihood of a promisor performing at least in part increases as we allow for continuous breach as opposed to just binary breach. Yet, the probability that a given contractual partner performs in full decreases. Performance rates are higher if a promise was received than if it was not received by the promisee.

Our paper casts some light on the old but now abandoned mutuality requirement in contracting, which held that "unless both are bound, neither is bound" (see Farnsworth 2004, §3.2),[1] and its remnant, the doctrine of consideration. We also gain insight on the doctrine of substantial performance and the doctrines of divisible obligations. Additionally, our paper suggests that we have to rely on more formal enforcement mechanisms, as the stakes of a transaction increase relative to income. This lends some support to the statute of frauds that requires contracts for the sale of goods over a certain threshold value to be in writing in order to be enforceable. Given that promisors do not feel as strongly bound by their promises if the promise was not received by the promisee, our results suggest that the common law mailbox rule does not track our moral intuition as well as its civil law counterpart, which requires the receipt of an offeree's acceptance.

[1] For an early expression of the mutuality requirement, see Harrison v. Cage, 87 Eng. Rep. 736 (K. B. 1698; either "all is a nudum pactum, or else the one promise is as good as the other").

The paper is organized as follows. Section 1 describes our design and procedure. Section 2 reports and discusses our results. Section 3 concludes.

1. Design and Procedure

1. 1. Design

In Experiment I, we let pairs of subjects play a dictator game. At the beginning of the game, a computer randomly selects a dictator and a recipient in each pair. The dictator then chooses between two actions, *Roll* or *Don't Roll*. If the dictator chooses *Don't Roll,* she receives $14, but the recipient receives nothing. If the dictator chooses *Roll*, she receives $10, and the recipient receives $12 with probability 5/6 and $0 with probability 1/6. Prior to the role assignment, each subject can send computer-based messages to "promise" or "not promise" to choose *Roll* if chosen as the dictator. The computer then randomly decides whether or not the message is delivered to the other side.[2]

Experiment II is almost identical to Experiment I, with two important modifications: First, in addition to the options *Roll* or *Don't Roll*, we introduce three intermediate *Roll* decisions: 1/4 *Roll*, 1/2 *Roll*, and 3/4 *Roll*. For the choices *Roll* and *Don't Roll,* the payoffs stay exactly the same. If the dictator chooses 1/4 *Roll*, she receives $13, and the recipient receives $3 with probability 5/6 and $0 with probability 1/6. If the dictator chooses 1/2 *Roll*, she receives $12, and the recipient receives $6 with probability 5/6 and $0 with probability 1/6. Finally, if the

[2] Except for the use of dollars rather than euros, this game is to a large extent identical to Vanberg (2008), who designed the dictator game to mimic the choice of the trustee in Charness and Dufwenberg's trust game (2006). What distinguishes our design from the other papers is that we allow for the possibility of messages not being delivered. We also use pre-coded messages as opposed to free-form communication.

dictator chooses 3/4 *Roll*, she receives \$11, and the recipient receives \$9 with probability 5/6 and \$0 with probability 1/6. The second major change from Experiment I is that the message of the recipient will never be delivered in Experiment II.

The first goal of the present paper is to decompose the effect of mutual promises. The design of previous papers was not suited to answer this question. Vanberg (2008) has the same experimental set-up as in Experiment I. However, in the communication phase he allows for free-form communication and therefore naturally ends up in a scenario where either all parties promise to each other or none of them promise. In other words, if there is a promise, there is also always mutuality. In Charness and Dufwenberg (2006, 2010), the agent promises to the principal, who then decides whether to entrust the agent with a joint project. Hence, an element of mutuality is established here also. We hypothesize that given the dictator chooses "promise," she is more likely to choose *Roll* if she received a promise from the pairing partner than if she received: i) a "no promise" message (H1a) or ii) no message from the pairing partner (H1b). This will cast light on the old but now abandoned mutuality requirement in contracting, which held that "unless both are bound, neither is bound" and its remnant, the doctrine of consideration.

The second goal of the paper is to study the dynamics of promising, *Roll* rates, and promise-keeping. The experiment consists of eight rounds. If mutual promising is important, we would expect that promising goes up. But there are also factors which might dampen promise-keeping over time. Subjects may see that they prefer choosing *Don't Roll* and might want to shun the experience of guilt due to breaking their promises. As Experiment I allows for mutual promises, whereas Experiment II never delivers the message of the recipient, we expect promising to increase in Experiment I (H2a) and to decrease in Experiment II (H2b), assuming that subjects learn over time how to maximize their payoff. If promising increases, we should wonder whether this will also cause more *Roll* decisions or if the increase in promising is purely strategic. We hypothesize that the law, by encouraging the exchange of commitments, contributes to the breeding of cooperation over time (H2c).

A third question is whether the introduction of the continuous breach decision in Experiment II affects contribution rates and contribution levels relative to Experiment I. We expect that the rate of contribution will increase as subjects who would not have contributed in a setting in which they faced a binary choice between contributing and not contributing might contribute at least a bit (H3a). But we also fear that the option of continuous breach of promises will decrease the level of contribution of those who already would have rolled in the binary setting (H3b). We hypothesize that the former effect will dominate the latter (H3c). This would be consistent with the doctrine of substantial performance and doctrines of divisible obligations.

A fourth goal is to investigate if promise-keeping depends on the dictator's income. We suspect that, given the same stakes, low income people are less likely to keep promises (H4). To the extent that the moral force of promise-keeping is a low cost enforcement mechanism acting as a substitute for more expensive formal enforcement, less promise-keeping by low income subjects would suggest a higher transaction cost for low income people given the same transaction value.

Finally, given the dictator chooses "promise," we ask whether she is more likely to choose *Roll* if her promise was received than if it was not received (H5). This could have two reasons: Either the fact that the message is not received weakens the commitment *per se*, or it reduces the guilt of the promise-breaking as a non-dictator who did not receive the message expected less. We try to disentangle these two motivations.

1.2. Procedure

We ran three sessions in February and March of 2012 with a total of 107 student subjects at the California Social Science Experimental Laboratory (CASSEL). The CASSEL subject pool consists of undergraduate students from UCLA. Subjects were assigned to visually isolated computer terminals. Beside each terminal they found paper instructions (reproduced in Appendix A). Questions were answered individually at subjects' seats (by the researchers).

We conducted two sessions of Experiment I in February and March with a total of 74 participants, 42 of whom were women. The average payoff in the first session was $15.30, including a $5 flat fee for arriving on time. The minimum payoff was $6.15, and the maximum payoff was $22. The average payoff in the second session was $20.65 (minimum: $11.00; maximum: $27.30). The higher payoff was due to a higher show-up fee of $10.[3]

We conducted one session of Experiment II in March with 33 participants, 14 of whom were women. Average payoffs were $15.50 (minimum: $5.50; maximum: $24.00). The show-up fee was $5.

Each session consisted of two practice rounds, for which subjects were not paid, followed by eight paying rounds. In each round, subjects interacted with another randomly chosen participant. Under no circumstances did participants interact with the same participant twice in the paying rounds. We achieved this by creating matching groups of at least ten participants. At the end of the experiment, one of the eight paying rounds was randomly chosen for payment (every round was

[3] We increased the show-up fee in the second session. Subjects had earned too little on average in the first session as the randomly selected round was one where the outcome of the rolling of the die was 1. Therefore all reciepients only got their show-up fee and modest payoffs from belief elicitation and a subsequent incentivized questionnaire.

equally likely). The amount paid out at the end of the experiment depended on the decisions made in that round. Elicitation of beliefs was incentivized. Subjects were paid for correct beliefs in all rounds except the one chosen for payment of the decision.[4] The 34 subjects in Session 1 received a fixed fee of $5 for arriving on time. The show-up fee was increased to $10 for the 40 subjects in Session 2. The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007).

The timeline of the experiment was as follows. First the subjects were randomly paired with an interaction partner. Then each subject chose one of two pre-coded messages: "If I will be in Role A [dictator], I promise to choose *Roll*," or "If I will be in Role A I do NOT promise to choose *Roll*."[5] The computer then randomly decided whether the message was delivered. The probability that each message went through was 50% and was independent of whether the pairing partner's message was received.[6] In either case, the sender of the message was informed whether the message was received by the other player or whether he or she received no message. During the communication phase, neither subject knew which member of the pair would subsequently be a dictator. In Experiment II, the dictator never received the message of the recipient, whereas the recipient could still get three different signals: "promise," "no promise," and "no message."

After the communication phase, one participant in each pair was randomly assigned to the role of a dictator or the role of a recipient. The role was randomly assigned anew in each round. It was always equally likely to be assigned to either role, regardless of the previous messages or actions in the game.

---

[4] To prevent hedging.

[5] In the instructions, we neutrally refer to the role of the dictator and the role of the recipient as "Role A" and "Role B," respectively.

[6] Therefore, subjects could not conclude from the fact that they did not receive their pairing partner's message, that their pairing partner did not receive their message.

Before the decision phase, all of the subjects saw a screen informing them of their role, reminding them of the message they had sent, and informing them whether this message was received. They also learned whether they received a message from their pairing partner and, if so, whether it was a promise. A screenshot of this screen is reproduced in the Appendix.

Next, dictator subjects were asked to submit their decisions, binary in Experiment I and continuous in Experiment II. At the same time, recipients were asked to guess their partners' decisions. Specifically, recipients were asked to choose a value from the five-point scale (see Figure 1). Each value is associated with payoffs that depend on the decision made by the partner. This procedure yields a five-point scale for first-order beliefs.

Subject A will	Certainly choose *Roll*	Probably choose *Roll*	Unsure	Probably choose *Don't Roll*	Certainly choose *Don't Roll*
Your Guess	o	o	o	o	o
Your earnings if the other player chooses *Roll*	\$0.65	\$0.60	\$0.50	\$0.35	\$0.15
Your earnings if the other player chooses *Don't Roll*	\$0.15	\$0.35	\$0.50	\$0.60	\$0.65

Figure 1: Guessing Payoffs

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 1 and asked to mark the box that they believed the recipient had clicked. If they guessed correctly, dictators earned $0.50. This yields a five-point scale for second-order beliefs.

At the end of each round, the dictator learned the payoff of both participants. In case the dictator chose *Roll*, she also found out the number that was rolled. The recipient learned only of his own payoff. Other than what can be concluded from this payoff, he did not learn which choice the dictator had made.[7] In Experiment II, the dictator further learned whether her second-order belief overestimated or underestimated the first-order belief of the recipient.

At the end of the experiment, we asked participants to fill out four computer-based questionnaires. The first questionnaire consisted of three incentivized questions measuring the tendency of our subjects to respond impulsively to cognitive tasks rather than after some reflection (see Frederick 2005).[8] The second questionnaire asked subjects to report in free form about the motivations behind their choices in the experiment. The third involved a "social desirability" questionnaire that measures the subject's tendency to reply to expectations in a manner that will be viewed favorably by others (see Fischer and Fick 1993). Finally, we asked a few demographic questions like gender, major, and income. Screenshots of the questionnaires can be found in the Appendix.

2. Results

2.1. Mutuality and Promise-Keeping

In Experiment I, a dictator faces one of three different scenarios when deciding whether to keep her promise: i) the dictator might have received a "promise" message from the recipient;

[7] We are interested in the moral force of promise-keeping. Besides anonymity, introducing some degree of deniability to a dictator who has chosen *Don't Roll* helps to isolate this effect.

[8] The participant could earn $0.50 per correctly answered question.

ii) she might have received a "no promise" message; iii) or she might not have received a message at all. Given the dictator had promised, receiving a "promise" message from the other side, as opposed to receiving a "no promise" message, increases *Roll* rates from 21% to 49%. The Wilcoxon rank-sum test based on matching-group-level averages of *Roll* decisions shows that this difference is highly significant (p = 0.01), supporting H1a. Interestingly, however, there is no significant difference in *Roll* rates between receiving a "promise" message and not receiving a message at all (49% vs. 45%; p = 0.44), showing no support for H1b. There is, however, a highly significant difference between receiving no message and receiving a "no promise" message (45% vs. 21%; p = 0.01). In other words, dictators do not reward recipients who have promised but instead punish those who have not promised (see Figure 2).

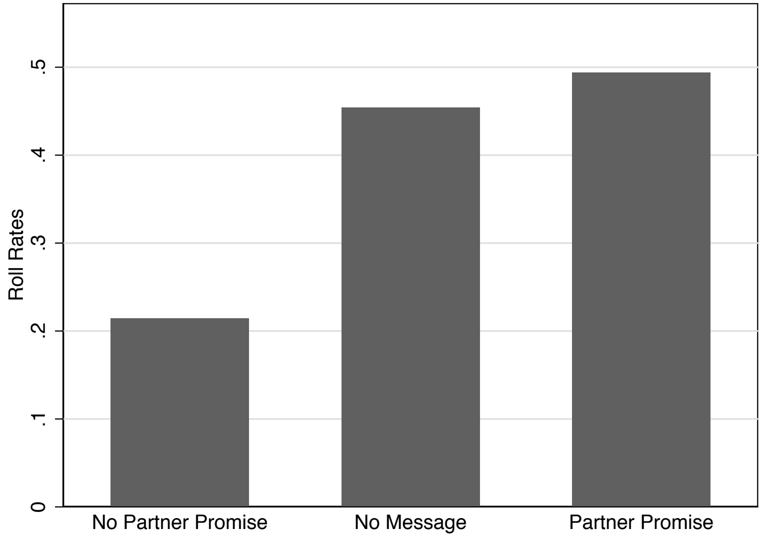


Figure 2: Promise-keeping depending on message received by the promisee.

This raises the question of whether our results support treating gratuitous promises differently from mutual promises. One could argue that our evidence does not support denying gratuitous promises' enforceability on the theory that the law should track our moral intuition. This is because promisors do not seem to require an affirmative return promise in exchange for their promises in order to feel bound by them. On the other hand, our results also suggest a big drop in promise-keeping if the promisor learns that his promisee affirmatively negated a return promise.

While the promisor seems to give the promisee the benefit of the doubt, this belief could very well be unstable. If a promisor can easily be led to believe that silence by the other side means refusal of a counter-promise, the erosion over time of the old principle that "unless both are bound, neither is bound" may make promise-keeping less stable, but studying this hypothesis is beyond the scope of this paper.

Our results might also offer an argument against means testing in social insurance systems if the state cares about the acceptance of the system by participants. If contributors would not receive any benefit from the system, acceptance would presumably decrease.[9]

2. 2. Dynamics of Promises and Promise-Keeping

The average rates of promising in Experiment I and Experiment II are very different (76% vs. 51%). The reason for that difference can be seen in Figure 3. While promise rates start out to be roughly the same, promising increases over time in Experiment I from 66% in period 1 to 82% in period 2. Regressing promises per round against the number of the round (clus-

[9] Participation in social insurance systems is not voluntary, but acceptance of the system by those who contribute to it is probably essential to its long-run viability.

tering standard errors on a matching group level) reveals that this relationship is highly significant ($p < 0.01$, Logit regression), supporting H2a. The effect is reversed in Experiment II (supporting H2b), where promise rates decrease from 61% in period 1 to 36% in period 8 ($p = 0.03$).

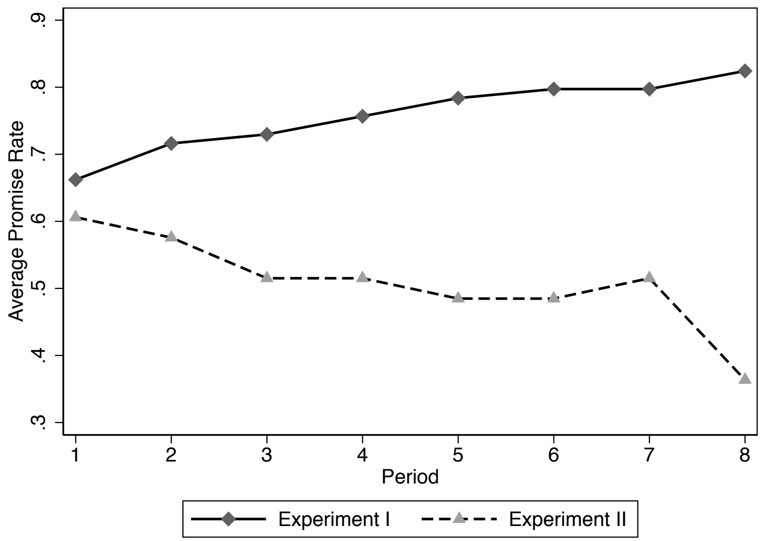


Figure 3: Promising over time when mutual promises can/cannot be exchanged.

There are two main differences between Experiment I and Experiment II. The first is that we made the *Roll* decision continuous. The second is that the dictator never received a message from her pairing partner. This second difference seems to be driving the dynamics of promising. In Experiment I, a dictator who has received a "no promise" message from the recipient is much less likely to *Roll*, as we have shown. Subjects may anticipate this reaction and might strategically promise in order to trigger positive reciprocity if the other party turns out to be the dictator. As the message of the non-dictator was never received in Experiment II, there is no such strategic advantage to promising in Experiment II. In Experiment II, a promisor only faces the cost of being bound if chosen as the dictator. Summing up, promise rates only increase if there is a strategic benefit to influence the other side and decrease otherwise.[10] Consistent with this, 36% of the 74 participants in Experiment I reported in a free-form questionnaire that they wanted to influence the other side with their promise, and 18% explicitly mentioned that they hoped their promise would lead the other party to *Roll*. In Experiment II, only 15% said they wanted to influence the other side, and only 6% said that they hoped for a positive reaction by the other party.[11]

This raises the question of whether increased promising in Experiment I also increased *Roll* rates over time. If this were the case, legal rules encouraging mutual promises would breed cooperation over time. But *Roll* rates seemed to decrease over time although the effect was far from significant ($p = 0.48$). Similarly, in Experiment II, the rate of contribution weighted with the level of contribution decreased over time, but also here the effect was not significant ($p = 0.26$).

Therefore, the possibility of mutual promises seems to increase promise-making over time but does not increase cooperative behavior. However, including the practice rounds, the decrease in cooperation over time became statistically significant for Experiment II ($p = 0.06$), while it remained far from significant for Experiment I ($p = 0.33$). Given that *Roll* rates decrease over time in both experiments, but that the effect is only significant in Experiment II (and only when we include the practice rounds), we find very weak support for the hy-

[10] This is an interesting methodological point for researchers who wish to study subjects' promise-keeping behavior. In such studies, researches will want to maximize the incidence of promise-making.

[11] These subjects had misunderstood the fact that, by design, their message was never received by the dictator in Experiment II.

pothesis that mutual promises might slow down the erosion of cooperation (H2c).

2.3. Continuous Performance Decision and Cooperation

In Experiment I dictators faced a binary choice between *Roll* and *Don't Roll*, whereas in Experiment II dictators made a more continuous choice between *Don't Roll*, 1/4 *Roll*, 1/2 *Roll*, 3/4 *Roll*, and *Roll*. How does moving toward a more continuous *Roll* decision affect the promise-keeping behavior of dictators?

In Experiment I, 44% of dictators who promised chose *Roll* (see Figure 4). In Experiment II, 62% of dictators who promised decided to forgo at least some payoff in order to give something to the other side. This increase in the contribution rate is statistically significant at the 10% level, supporting H3a ($p = 0.09$, two-sided ttest run on matching group averages).[12] However, the contribution rate weighted with contribution levels was 46% and, therefore, only slightly higher than in Experiment I. Moreover, this difference is not statistically significant ($p = 0.85$), offering no support for H3c. In other words, people are more likely to give something if they face a continuous *Roll* decision, but when they give, they give less (supporting H3b).

In contrast, given no promise, the percentage of people who contribute in Experiment II (9.1%) barely increases compared to Experiment I (8.6%), and the weighted contribution rate in Experiment II even goes down (2.7%, see Figure 4).[13]

---

[12] Running the Wilcoxon ranksum test with matching-group-level averages gives us $p = 0.13$, and $p = 0.09$ if we include the practice periods.

[13] 8.6% vs 9.1%, two-sided t-test, $p = 0.80$; 8.6% v 2.7%, two-sided t-test, $p = 0.23$.

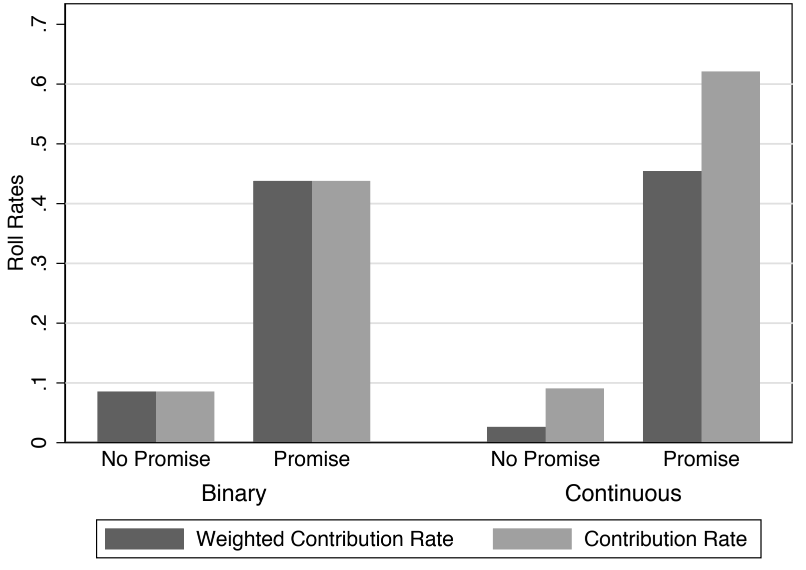


Figure 4: Rates and Levels of Performance under Binary and Continuous Breach.

Although the difference in *Roll* rates between the conditions in which the dictator promises and those in which she does not promise is statistically significant, we cannot conclude that promising causes higher *Roll* rates. People who are more likely to *Roll* also could be more likely to promise, and we picked up a mere sorting effect. However, Charness and Dufwenberg (2006) have shown that the opportunity to communicate increases *Roll* rates compared to a setting where communication is impossible.[14]

[14] Charness and Dufwenberg (2010) could not replicate this result with pre-coded messages, but Stone and Stremitzer (2020) were able to do so in a similar setting.

Still, the difference between results in Experiment I and II casts some light on the effect of legal rules by comparing a setting where the breach decision is binary to a situation where the breach decision is continuous. The results suggest that the effect of the doctrine of substantial performance or of doctrines allowing for the divisibility of obligations will increase the likelihood of contractual partners performing at least in part but will also decrease the probability that a given contractual partner will perform in full. The main effect of those doctrines may, therefore, lie more in reducing the variance in performance than in increasing the overall level of performance.

2.4. Income and Promise-Keeping

In this section we investigate how the rate of promise-keeping depends on income. In order to create a measure of income, we asked subjects in a post-experiment questionnaire to compare their parents' income to the parents' income of other UCLA students. We found in Experiment I that 77% of promises made by subjects who reported their family to be "Affluent" were kept, while only 6% of promises were kept by subjects who reported their family to be "Very Poor" (see Figure 5). The Wilcoxon ranksum test, run on subject averages, shows that differences in promise-keeping between income categories are for the most part statistically significant (supporting H4).[15]

---

[15] Of all possible six comparisons between income categories, only the Affluent vs. Average (p = 0.20) and the Poor vs. Very Poor (p = 0.52) comparisons are not statistically significant. All other comparisons are significant at at least the 10% level. We base our analysis on subject averages, as this seems to be the most natural way given we are interested in the effect of a subject's income level.

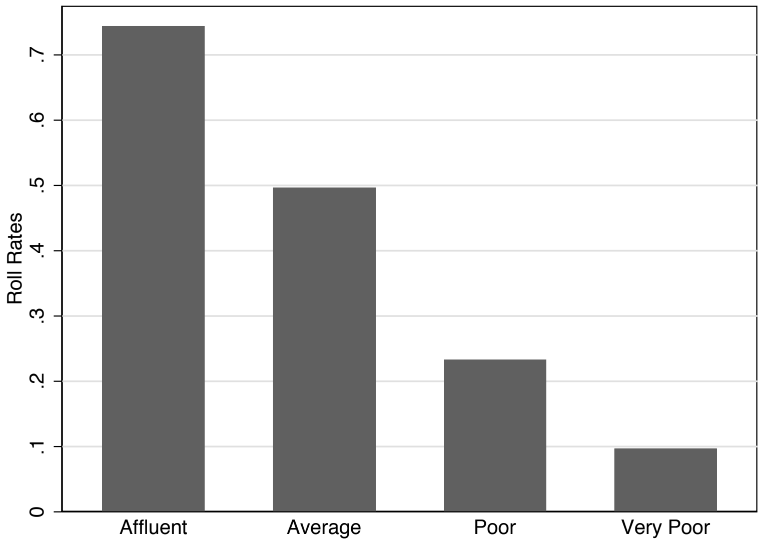


Figure 5: Income and Promise-Keeping.

Experiment II further supports this result. The decrease in exact promise-keeping, that is, the decrease in choosing *Roll* is most pronounced, followed by the decrease in the weighted rate of contribution, and the decrease in the rate of contribution.[16]

In Experiment II, in addition to asking subjects about their parents' income, we also asked them about their monthly spendable income. In a linear regression model, we find that the positive relationship between income and the weighted rate of contribution is significant at the 5% level.[17] If we drop

[16] The result is mainly driven by the difference between those who self-report their family's situation as "Affluent" and those who self-report their family's situation as "Average."

[17] Significance disappears if we cluster standard errors at the subject level (p = 0.19). However, all other results mentioned in this paragraph hold, irrespective of whether we cluster at the subject level or not.

those observations where spendable income is zero, we get significance at the 1% level, and the R^2 increases from 6% to 34%. Results do not change if we drop the outlier (monthly spendable income: $2,000) or if we control for cognitive reflection scores or social desirability scores.

To the extent that we accept social desirability scores as a measure of pro-social attitude, it is interesting that, if anything, there is a weak negative relationship between income and social desirability scores.

We can conclude that promise-keeping increases with parents' income and the participant's spendable income. This could not have been due to a generally higher pro-social attitude as picked up by the social desirability score. Therefore, the effect may have been due to either: i) the higher relative cost of being "moral" for lower income participants or ii) different kinds of "morality," with promise-keeping being a very special kind of morality not picked up in the social desirability score.

To the extent that the legal system, with its extrinsic mechanism of promise enforcement, is a substitute for promise-keeping as driven by moral considerations, this result suggests that for transactions of similar absolute value, promisees of lower income promisors need to rely on formal enforcement more than promisees of higher income promisors. This creates additional transaction costs when dealing with lower income promisors. The result also offers support to the statute of fraud provisions kicking in for higher-stake transactions. [18] These findings are consistent with that of Ederer and Stremitzer (2017), who offer suggestive evidence from a structural estimation that about one quarter of subjects in their sample trade off promise-keeping and money.

[18] The statute of fraud requires a promise involving the sale of goods to be in writing in order to be enforceable when the value of the promise exceeds a certain threshold. Thus the statute of fraud facilitates formal enforcement in cases in which it is more likely to be required.

2. 5. Receipt of Promise and Promise-Keeping

Does a promisor feel less bound by her promise if the promise was not received by the promisee? In Experiment I, we find that with a binary *Roll* decision, *Roll* rates decrease from 46% to 41% in the aggregate when comparing cases in which the promise was received by the promisee to those in which it was not received. This effect is not significant (p = 0.57, Wilcoxon ranksum).

In Experiment II, where the *Roll* decision is continuous, the difference in weighted contribution rates between the *receive* and the *non-receive* condition is stronger (52% vs. 39%), but the difference is not statistically significant (p = 0.15). The difference is even higher when comparing the contribution rate (73% vs. 53%). This difference in contribution rates is significant, confirming H5 (p = 0.04, Wilcoxon ranksum test).

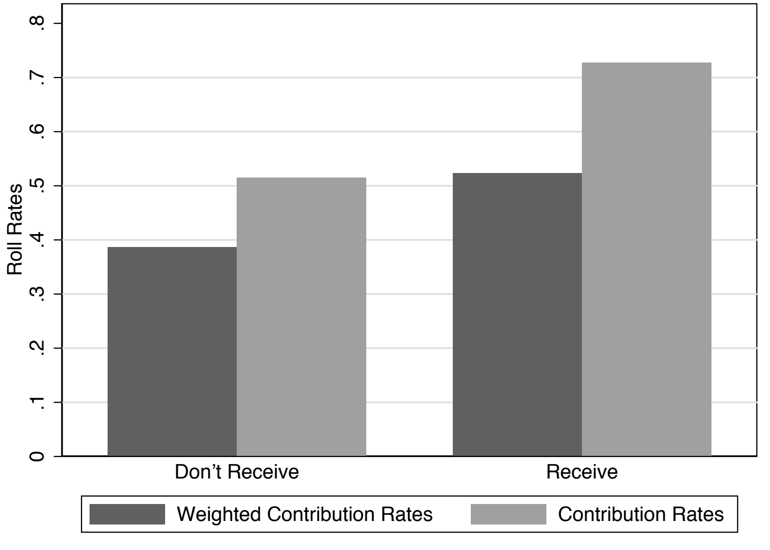


Figure 6: Receipt of promise and promise-keeping.

On the theory that law should track our moral intuitions, our results suggest that we should be suspicious of the common law "mailbox rule," which holds offerees to their promises as soon as they have dispatched their acceptance, that is, before it is received by the other party. Participants in our experiments felt less bound by a promise that was not received than by a promise that was received. The common law "mailbox rule," therefore, departs from our moral intuitions further than its civil law counterpart, which requires the other party to receive the promise in order for it to be binding.

The observed difference could have two potential reasons. First, it could be that the promisor feels less guilt to let down somebody who has not received the promise. This is because a potential promisee who did not receive a message has a lower level of expectation in performance than somebody who has received a promise. After all, the promisee cannot be sure of whether or not a promise has been given to him. This would be a guilt aversion theory of promise-keeping (Charness and Dufwenberg 2006). A second explanation could be that the promisor does feel less committed to her promise *per se* (see Vanberg 2008). The promisor may feel that she is not bound at all if the other party has not received her promise. That is, promising, at least for some promisors, may be bilateral in nature, requiring the promisee's notice as opposed to a vow somebody secretly makes to herself.

2. 6. Mechanism

As we explained, it is difficult in our design to disentangle two possible explanations for promise-keeping. This is because, given a promise, both expectations and commitment *per se* are influenced by the *receive/non-receive* manipulation. However, one potential way of investigating whether aversion to disappoint a recipient's expectations plays a role in promise-keeping is to look at whether recipients' expectations matter in cases where no promise was given by the dictator. The level of commitment *per se* is zero, and therefore constant, independent of whether the message was received. The only thing that changes is the nature of the promisee's expectations. This is because the recipient plausibly has higher expectations that the dictator will choose *Roll* when he has received no message than when he has received the "no promise" message. In the former case, the recipient can reasonably expect that the dictator promised with a certain probability. However, in the latter case, the recipient knows for sure that the dictator has not promised. Therefore, the *receive/non-receive* manipulation allows us to vary first- and second-order expectations while keeping commitment *per se* at zero.

In Experiment I, we find a small effect in the right direction: 6% vs. 11% of *Roll* choices depending on whether the message was received or not received. However, there are very few observations, and the effect is not significant ($p = 0.23$, Wilcoxon ranksum test). There is a stronger effect for women (11% vs. 20%), which is almost significant at the 10% level ($p = 0.12$).[19] If the effect is present, we would expect it to be more visible in Experiment II where we made promise-keeping continuous. Yet, the effect is not significant and even goes in the opposite direction ($p = 0.32$). We also find no support for any gender effect in Experiment II.

If we had found an effect on *Roll* rates, we could conclude that recipients' expectations matter. However, we only find a small, non-significant effect in the right direction in Experiment I and a non-significant effect in the wrong direction in Experiment II. The fact that expectations do not seem to matter in the absence of promising is in line with results by Vanberg (2008) and Ellingsen et al. (2010), who find that expectations

[19] There are only six observations where subjects chose *Roll* without promising. All of the subjects were women. In four of the six cases, the message was not received.

are not independently significant. Our finding, however, runs counter to the theory by Charness and Dufwenberg (2006) and evidence from studies by Reuben et al. (2009), Bellemare et al. (2011), and Regner and Harth (2014). As Mischkowski et al. (2019) show, and as Ederer and Stremitzer (2017) had hypothesized, these conflicting results in the literature can be reconciled by accounting for one important fact: the expectation *per se* effect, that is, the effect of expectations which are unsupported by a promise, is much weaker than the interaction effect, that is, the effect of expectations which are supported by a promise.[20] Hence, studies sometimes seem to pick up the guilt aversion, and sometimes they do not.

We also find that mutuality does not matter in either experiment in the no-promise condition. Accepting that the law should track our moral intuitions, the result that neither the recipient's expectations nor the mutuality of promises matter in the absence of a promise by the dictator lends support to the very cautious way in which the law accepts claims in unjust enrichment.[21]

3. Conclusion

The goal of this paper was to study lay people's moral intuitions of promise-keeping as a way to inform us of the extent to which these intuitions correspond to legal doctrines in which the concept of a promise is central. We find that a promisor who—because of a technical glitch —did not receive a return promise is not less likely to keep her promise than a promisor who received a return promise from the promisee. This seems to suggest that treating mutuality as a requirement to enforce promises does not correspond to our moral intuitions. Howev-

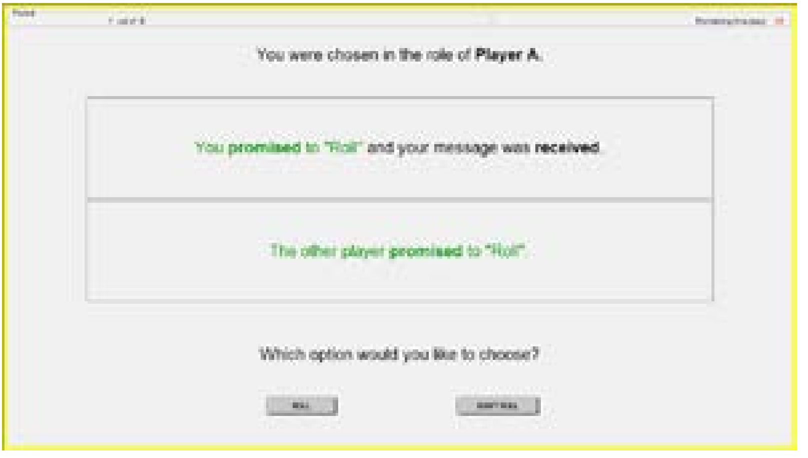
---

[20] Ederer and Stremitzer (2017) find that expectations that are backed up by a promise matter for promise-keeping.

[21] Unjust enrichment is a doctrine that recognizes that a person may incur an obligation without that person having given a promise.

er, we also find that a promisor feels much less bound by her promise if the promisee explicitly declines to make a return promise. Therefore, it seems that a promisor who does not receive a message gives the promisee the benefit of the doubt. But this belief might be unstable, and therefore, the mutuality requirement might still be important, as it might serve the function of stabilizing the promisor's beliefs about the promisee's willingness to give a return promise. Where mutual promises could be exchanged, we also saw that participants tried to strategically trigger positive reciprocity by promising, a motive different from the more commonly recognized motive to promise in order to trigger a return promise. Another result suggests that allowing promisors to choose intermediate levels of performance, as opposed to forcing them into a binary decision whether or not to keep their promises, has the effect of making promisors more likely to cooperate partially but less likely to cooperate in full. Indeed, the average cooperation level does not increase. Therefore, making the performance decision more continuous, as promoted by the doctrine of substantial performance and divisibility, mainly seems to reduce the variance in performance rather than increase the average level of performance. We also find that promisors who self-report lower income are less likely to keep promises. However, income is not positively correlated with pro-social behavior. This would be consistent with promisors trading off the marginal value of money with the guilt experienced from breaking a promise. One implication of this finding is that transaction costs become higher as stakes of a transaction increase relative to the income of the promisor as parties might need to rely more on formal enforcement of promises. Finally, we find that promisors do not feel as strongly bound by their promises if the promisee did not receive the promise. Our results suggest that the common law mailbox rule does not track our moral intuitions as well as its civil law counterpart, which requires the receipt of an offeree's acceptance.

4. Appendix


You were chosen in the role of **Player A**.
You **promised** to "Roll" and your message was **received**.
The other player **promised** to "Roll".
Which option would you like to choose?


Demographics

Acknowledgements

The authors are grateful to Jennifer Arlen, Seana Shiffrin, Steve Lippman and seminar audiences at UCLA and NYU. The authors also thank Juan Armas Pizzani and Henry Kim for their help in preparing the manuscript.

Bibliography

Bellemare, Charles, Alexander Sebald, and Martin Strobel. 2011. “Measuring the Willingness to Pay to Avoid Guilt: Estimation Using Equilibrium and Stated Belief Models.” *Journal of Applied Econometrics* 26 (3): 437–453.

Charness, Gary, and Martin Dufwenberg. 2006. “Promises and Partnership.” *Econometrica* 74 (6): 1579–1601.

Charness, Gary, and Martin Dufwenberg. 2010. “Bare Promises: An Experiment.” *Economics Letters* 107 (2): 281–283.

Che, Yeon-Koo, and Donald B. Hausch. 1999. “Cooperative Investments and the Value of Contracting.” *American Economic Review* 89 (1): 125–147.

Ederer, Florian, and Alexander Stremitzer. 2017. "Promises and Expectations." *Games and Economic Behavior* 106: 161–178.

Ellingsen, Tore, Magnus Johannesson, Sigve Tjøtta, and Gaute Torsvik. 2010. "Testing Guilt Aversion." *Games and Economic Behavior* 68 (1): 95–107.

Farnsworth, E. Allan. 2004. *Contracts*. 4th ed. New York: Aspen Publisher.

Fischbacher, Urs. 2007. "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2): 171–178.

Fischer, Donald G., and Carol Fick. 1993. "Measuring Social Desirability: Short Forms of the Marlowe-Crowne Social Desirability Scale." *Educational and Psychological Measurement* 53 (2): 417–424.

Frederick, Shane. 2005. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives* 19 (4): 25–42.

Mischkowski, Dorothee, Rebecca Stone, and Alexander Stremitzer. 2019. "Promises, Expectations, and Social Cooperation." *Journal of Law & Economics*, forthcoming.

Myerson, Roger B., and Mark A. Satterthwaite. 1983. "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory* 29 (2): 265–281.

Regner, Tobias, and Nicole S. Harth. 2014. "Testing Belief-Dependent Models." *Max-Planck Institute of Economics Jena Working Paper*.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2009. "Is Mistrust Self-Fulfilling?" *Economics Letters* 104 (2): 89–91.

Shiffrin, Seana Valentine. 2008. "Promising, Intimate Relationships, and Conventionalism." *Philosophical Review* 117 (4): 481–524.

Stone, Rebecca, and Alexander Stremitzer. 2020. "Promises, Reliance, and Psychological Lock-In." *Journal of Legal Studies*, forthcoming.

Vanberg, Christoph. 2008. "Why Do People Keep Their Promises? An Experimental Test of Two Explanations." *Econometrica* 76 (6):1467–1480.

Florian Ederer
Yale School of Management
florian.ederer@yale.edu

Alexander Stremitzer
Center for Law & Economics
ETH Zurich
astremitzer@ethz.ch