Games and Economic Behavior 113 (2019) 137–146

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Promises, expectations & causation ☆

Giovanni Di Bartolomeo [a], Martin Dufwenberg [b,c,d], Stefano Papa [a], Francesco Passarelli [e,f,*]

[a] *Sapienza University of Rome, Italy*
[b] *University of Arizona, United States of America*
[c] *University of Gothenburg, Sweden*
[d] *CESifo, Germany*
[e] *University of Turin, Italy*
[f] *Bocconi University, Milan, Italy*

ARTICLE INFO

Article history:
Received 3 November 2017
Available online 9 August 2018

JEL classification:
A13
C91
D03
D64

Keywords:
Promises
Expectations
Guilt aversion
Moral commitment
Causation

ABSTRACT

Why do people keep their promises? Vanberg (2008) and Ederer and Stremitzer (2017) provide causal evidence in favor of, respectively, an intrinsic preference for keeping one's word and Charness and Dufwenberg's (2006) expectations-based account based on guilt aversion. The overall picture is incomplete though, as no study disentangles effects in a design that provides exogenous variation of both (the key features of) promises and beliefs. We report evidence from an experimental design that does so.

1. Introduction

Promises often foster trust and cooperation. A recent literature explores why. Two explanations have been proposed, and experimental tests provided some support and some controversy:

☆ We thank Giuseppe Attanasi, Pierpaolo Battigalli, Gary Charness, Giuseppe Ciccarone, Fredrik Carlsson, Christoph Vanberg, two anonymous referees, an Associate Editor, and participants in the 2nd BEG Workshop on "Behavioural and Experimental Economics" at the University of Gothenburg and in the workshop on "Communication, Reciprocity and Beliefs" at Sapienza University of Rome for their very useful comments. Francesco Passarelli gratefully acknowledges financial support of Baffi-Carefin Center and Collegio Carlo Alberto.
* Corresponding author.
E-mail addresses: giovanni.dibartolomeo@uniroma1.it (G. Di Bartolomeo), martind@eller.arizona.edu (M. Dufwenberg), stefano.papa@uniroma1.it (S. Papa), francesco.passarelli@unito.it (F. Passarelli).

https://doi.org/10.1016/j.geb.2018.07.009
0899-8256/

- Charness and Dufwenberg (2006) (C&D) propose an *expectation-based explanation* (EBE). People are guilt averse (GA): person i feels bad if he hurts j relative to j's expectations.[1] A promise from i to j, then, changes j's expectation and i does not renege because he would feel guilty if he did.[2]
- A *commitment-based explanation* (CBE), popular in social psychology, posits that people have a preference for promise keeping *per se*. Vanberg (2008) proposes a novel methodology to test it.[3]

EBE thus combines GA with the idea (not implied by GA) that promises shape beliefs. CBE instead posits that promises have a direct causal effect on promise keeping because people have an intrinsic motivation to keep their word. The empirical implications of these two theories are substantially different. CBE implies that people are likely to honor a promise once they have given their word. EBE predicts that they will keep their promise only if it raised the counterparts' expectations. In principle, however, the two explanations are not mutually exclusive. It is plausible that an individual has some inclination both to keep her word and to live up to the higher expectations that her promise may have created.

In C&D's experiment messages are observed and beliefs elicited. This is central to their tests, but the conclusions rely on home-grown variation of messages and beliefs across subjects. Herein lie two potential problems. First, since subjects are randomly assigned neither to their messages nor to their beliefs, C&D's result may reflect correlation rather than causation. Second, as a result of the first problem, C&D cannot independently evaluate EBE and CBE.

Vanberg (who coined the terms EBE and CBE) focuses on the second problem, arguing that C&D's results are confounded in that CBE rather than EBE may be the driving force. He proposes an ingenious "partner-switching" design (described below) which induces exogenous variation in whether a player who sent a promise is actually paired with whoever received that promise. He is thus able to test CBE as a cause of promise-keeping, and reports support. Vanberg did not address the empirical relevance of EBE though. He does not report tests analogous to C&D, regarding within-treatment correlations of choices and beliefs of players who have to decide whether to keep their promise or not. Even if he did, such data would in part suffer from the same problem as in C&D: variation in expectations would not be exogenously induced.[4]

Ederer and Stremitzer (2017) (E&S) propose a novel design, including an "unreliable random device" (see footnote 6), which creates exogenous variation in players' expectations. They are able to test EBE as a cause of promise keeping, and report support. E&S do not address the empirical relevance of CBE though. Their design includes no analog to Vanberg's partner-switching feature.

To sum things up, Vanberg and E&S provide evidence supporting, respectively, CBE and EBE. The overall picture is incomplete, however, since neither study disentangles CBE and EBE in a design that provides exogenous variation of *both* promises and beliefs. Our main goal is to fill this gap. We report experimental evidence from a design with exogenous variation of promises as well as of beliefs, allowing a test of the relative importance of EBE vs. CBE. The main hypothesis to be tested is whether variation in second-order beliefs matters for promise-keeping, and if so how the magnitude of that *let-down* effect compares to the commitment effect.

Section 2 describes design and procedures, Section 3 proposes a check test for validity of design, Section 4 presents our main results, and Section 5 sums up.

2. Experimental design

We use Vanberg's binary-choice random-dictator game with a partner-switching mechanism. To explain the rules precisely, it is helpful to first consider the game depicted in Fig. 1: a binary-choice random-dictator game *without* partner-switching.

It is randomly determined whether player 1 or 2 will be the "dictator" who chooses between *Don't Roll* and *Roll*. The other player will be the "recipient." Payoffs are as indicated.[5]

Vanberg's design makes two critical changes relative to Fig. 1. First, before subjects are told their roles, they are given the opportunity to communicate and (if they so choose) exchange promises. Then, they are randomly assigned their roles, dictator or recipient. This is when the second change appears. To generate exogenous variation in promises, *some of the*

---

1 Battigalli and Dufwenberg (2007) present a general model of GA. Many papers report experimental evidence supporting GA in the absence of promises (e.g., Dufwenberg and Gneezy, 2000; Guerra and Zizzo, 2004; Bacharach et al., 2007; Reuben et al., 2009; Chang et al., 2011; Bellemare et al., 2011; Bracht and Regner, 2013; Regner and Harth, 2014; Khalmetski et al., 2015). Ellingsen et al. (2010) do not find support, but Khalmetski et al. (2015) offer a possible reconciliation.

2 At a conceptual level, it is useful to clarify the difference between EBE and GA, as the former requires the latter but not vice versa. GA captures a general aversion to disappoint the expectations of others, whereas EBE captures the individual's aversion to disappoint expectations that may be somehow induced by a promise.

3 See Braver (1995), Ostrom et al. (1992), Ellingsen and Johannesson (2004), and Vanberg (2008) for exploring related notions. C&D too discuss the idea (Section 5.2) but argue against it. Krupka et al. (2017, p. 1708) argue, and study experimentally, that promise-keeping behavior may also be determined by the desire to comply with social norms. They are interpreted as social constructions that define how individuals ought to behave in specific situations: individuals then experience disutility when they take actions that are collectively judged inappropriate. See also Bicchieri (2006).

4 Vanberg (2008) obtains exogenous variation in beliefs, but only for players whose partner has been *switched* (i.e., players who face a partner that eventually received a promise from someone else). These beliefs cannot be used to test EBE because, to do it, one needs to induce exogenous variation in the beliefs of *non-switched* players *who made* a promise (i.e., players who face a partner that received the promise from them).

5 Payoffs and names of choices in subgames match those found in C&D's related trust game.

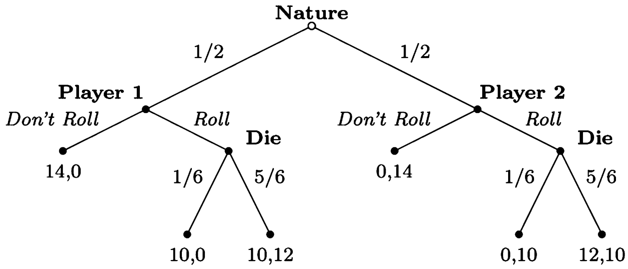

Fig. 1. A binary-choice random-dictator game.

recipients are randomly re-matched with a new dictator, according to a given switching probability (see below). Each dictator whose co-player was switched then can read the messages that occurred between the new recipient and the dictator with whom that new recipient was initially matched. Finally, each dictator (whether involved or not in a switch) chooses between *Roll* and *Don't Roll*, like in Fig. 1.

The design achieves exogenous variation in whether a player who sent a promise is paired with someone who received that promise. It involves asymmetric information between dictators and recipients; all subjects know that they will be re-matched with a certain probability, but after re-matching only dictators are told whether their recipient was switched. Therefore, recipients' first-order beliefs (and in turn dictators' second-order beliefs) can depend on whether or not they received a promise, but not on whether there was a switch.

In Vanberg's study, the switching probability is always 50%. To achieve exogenous variation not only regarding promises but also regarding expectations, we crucially introduce a random change in the switching probability across treatments[6]: either *high* (75%) or *low* (25%). Recipients' first-order beliefs, and dictators' second-order beliefs, at the time the dictator makes her choice, may now plausibly and directionally depend on the value of the switching probability. Namely, in light of the relevance of CBE (as documented by Vanberg), it is plausible that people expect dictators to be more inclined to keep their own promise than a promise made by someone else. Hence, recipients who received a promise should expect it to be kept with higher probability if the switching probability is low (i.e., 25%) rather than high (i.e. 75%). And, if dictators understand that, their second-order beliefs should vary by switching probability in the same direction as recipients' first-order beliefs.

This explains how our treatment variable, the switching probability, achieves exogenous variations in first- and second-order beliefs. We use that feature to test the effect of second-order beliefs on promise-keeping. A higher *Roll* rate in the presence of higher second-order expectations would support that the dictators' behavior is causally affected by their expectations, as predicted by EBE. Moreover, following Vanberg, we use exogenous variation in promises to test CBE. A higher *Roll* rate in the cohort of non-switched dictators would support that people have a stronger taste to keep their own rather than others' promises. Then, as predicted by CBE, making a promise *per se* would determine promise-keeping behavior.

Our tests for EBE and CBE are thus based on exogenous variations in expectations and promises. They can be summarized as follows:

H1: EBE. We consider the non-switched dictators who made a promise. We compare the average behavior in the group with high second-order expectations to the average behavior in the group with low second-order expectations (as induced by low vs. high switching probability). A higher average *Roll* rate in the first group would support EBE.

H2: CBE. Holding constant the switching probability (and thus holding dictators' second-order beliefs constant), we compare the average behavior of non-switched dictators with the average behavior of switched dictators that are re-matched with a co-player who received a promise from someone else. A higher average *Roll* rate in the cohort of non-switched dictators would support CBE.

While EBE refers to the behavior of *non-switched* dictators, our experimental data also enable us to test the GA of *switched* dictators. GA tests can be obtained from two different kinds of exogenous variation in expectations. The first one is due to the change in the switching probability (as in the test of EBE). The second one is due to prior communication. Recall that switched dictators can read prior communication of their re-matched co-player. Thus, the second-order beliefs of dictators that are re-matched with recipients who received a promise in their previous communication are higher than the second-order beliefs of dictators that are re-matched with recipients who received no promise. Hence, we perform two tests of GA.

H3(a): GA. We consider switched dictators that are re-matched with a subject who received a promise from someone else. These dictators can be split in two groups, according to the value of the switching probability. Second-order beliefs are higher in the group with low switching probability. A higher average *Roll* rate in that group would support GA.

[6] The idea to use the partner-switch probability as a treatment variable is inspired by E&S. Their design is otherwise different from ours. Instead of using Vanberg's game, they introduce a modified version of the C&D's trust game. Instead of a switching probability, they have a move by Nature determining whether the trustee will have a choice. The associated probability can be high or low, and only the trustee observes Nature's choice; this too yields exogenous belief-variation.

H3(b): GA. Holding constant the switching probability, we compare the group of switched dictators facing recipients who received a promise to the group of switched dictators facing recipients who received no promise. Second-order beliefs are higher in the first group. A higher average *Roll* rate in the first group would support GA.

The experiment was conducted at the *CIMEO Experimental Economics Lab* of Sapienza University of Rome (December 2015).[7] The design involved 192 undergraduate student subjects (6 sessions, 32 subjects each), recruited using an online system. Upon arrival, subjects were randomly assigned to 32 isolated computer terminals.[8] Three assistants handed out instructions (cf. Supplementary Material online) and checked that participants correctly followed the procedures. Before playing any game, subjects filled out a short questionnaire testing their comprehension.

Each session consisted of eight rounds, with perfect stranger matching. Payoffs, as shown in Fig. 1, were computed in tokens (where 1 token = 0.5 euro). At the end of each session, one round was randomly chosen for payment. First- and second-order beliefs were elicited by asking subjects to guess their counterparts' actions and guesses, respectively. Incentives were provided for all rounds except the one chosen for payment, implying that subjects had no incentive to hedge against bad outcomes and thus to misreport their beliefs.[9] All subjects received a fixed show-up fee of 2.50 tokens.

Each round implemented the following sequence of five stages:

1. **Communication.** Subjects were randomly matched to form 16 chatting pairs, with random determination of who would start the chat. As in Vanberg's design, each chat consisted of four one-way messages in sequence. Each message could be of at most 90 characters, and was cataloged as involving a promise or not (see below).
2. **Role assignment and revelation of the switching probability.** After communication, roles were randomly assigned in each pair and were subjects informed of that. Depending on treatment, the switching probability was announced as either 25% or 75%.
3. **Belief elicitation.** This stage has two parts:
   a. First-order beliefs: each recipient was asked to guess about his/her payoff.
   b. Second-order beliefs: dictators were asked to guess the guess of the person with whom they had formed a chatting pair.
4. **Switching.** Some recipients were switched: 25% or 75%, depending on treatment. Only dictators were informed whether a switch occurred. Dictators with switched recipients were then allowed to read the prior conversation of their new recipients.
5. **Dictators' action.** All dictators made their choice: *Roll* or *Don't Roll*. All subjects were informed of their payoff for the round. Recipients were not informed whether they had been switched or not, nor could they infer the dictator's choice when their payoffs were zero.[10]

Following C&D (2006), we elicit beliefs before dictators make their choice. Since we wish to test the effect of expectations on the behavior of *non-switched* dictators who sent a promise, we elicit their beliefs before they know whether their co-player has been switched or not. Exogenous variation in their expectations can only be due to a change in the switching probability. In this respect we differ from Vanberg (2008), who elicits beliefs after dictators are switched and after they make their choice.

As mentioned earlier (footnote 4), in testing CBE, Vanberg does not induce exogenous variation in the beliefs of non-switched dictators. Therefore, he cannot test the effects of SOBs on own promise keeping. However, he obtains one case that is useful for a test of GA. Comparing switched dictators matched with recipients who previously received a promise to those matched with recipients who did not receive it, he obtains an exogenous variation in some switched dictators' beliefs. Looking at all dictators, we can also perform a test for GA (H3(a)). However, since we elicit beliefs before any switch occurs, we do not observe the beliefs of switched dictators, but we can assume that the exogenous variation in expectations due to a change in the switching probability also holds for them.[11] This is reasonable since dictators know that recipients do not observe whether there is any switch, hence their first-order beliefs should depend on the switching probability and not on whether a switch occurred.

Our sample consists of 768 dictator decisions, equally split between switch and no-switch subsets. Messages were classified according to Vanberg's protocol; each pair of messages sent by a subject in a round was treated as a unit. Hence, we had 1,536 messages. We asked a research assistant to code these according to whether they conveyed a promise or a statement of intent stating that the subject would *Roll*, thus obtaining 575 promises out of 768 messages.[12]

---

[7] This is the acronym for *Center for Interpretation and Modeling of Experimental Observations*.

[8] The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

[9] Our elicitation procedure is described in detail in Appendix A.

[10] Recipients could obtain a zero payoff in two cases: (i) their dictator had chosen *Don't Roll*; (ii) their dictator had chosen *Roll* and the outcome of the die-roll was "1".

[11] Recall that we elicit beliefs of all dictators before switch takes place. We do not elicit beliefs of *switched* dictators regarding the beliefs of their *re-matched* recipients.

[12] We use the same coding strategy as Vanberg, treating each couple of messages sent by any subject in any round as one single message, and that message was evaluated by a research assistant. It is worth noting that the frequency of promises does not statistically differ across high/low switching probability treatments (i.e., 76% vs. 73%; $Z = 1.16$, $p = 0.247$).

Table 1
First-order and second-order beliefs (768 obs.).

Switching probability	Promise (575 obs.)		No promise (193 obs.)	
	1st-order beliefs (a)	2nd-order beliefs* (b)	1st-order beliefs (c)	2nd-order beliefs* (d)
Low (25%)	0.74 (s.d. 0.30, obs. 282)	0.80 (s.d. 0.27, obs. 282)	0.42 (s.d. 0.38, obs. 102)	0.38 (s.d. 0.38, obs. 102)
High (75%)	0.55 (s.d. 0.34, obs. 293)	0.66 (s.d. 0.33, obs. 293)	0.41 (s.d. 0.32, obs. 91)	0.40 (s.d. 0.35, obs. 91)

* Second-order beliefs are elicited for all dictators, before they know whether they have been re-matched or not.

Table 2
Testing EBE (286 obs.).

Switching probability	Roll rates	Beliefs	
		1st-order beliefs	2nd-order beliefs
	(a)	(b)	(c)
(1) Low (25%)	51% (s.d. 0.50, obs. 214)	0.74 (s.d. 0.30, obs. 214)	0.80 (s.d. 0.27, obs. 214)
(2) High (75%)	58% (s.d. 0.50, obs. 72)	0.51 (s.d. 0.35, obs. 72)	0.65 (s.d. 0.35, obs. 72)

3. Expectations and switching probabilities

Before we explore the causal effect of promises and second-order beliefs, we must establish that our design provides an adequate test bed in the sense of inducing exogenous variation in first- and second-order beliefs. Table 1 presents those beliefs, in the form of fractions of dictators who chose to roll.[13] It also reports standard deviations (s.d.) and observations (obs.). All reported statistics adopt the Wilcoxon signed rank test, which compares averages at the session level. Our data are independent at session level, but not at the individual level. The Wilcoxon signed rank tests accounts for such structure in the data. We check the robustness of our results by using also individual data. Both Wilcoxon rank sum test and Fligner–Policello test confirm the Wilcoxon signed rank test, so we omit to report them.

Our treatment variable of different switching probabilities is intended to create exogenous variation in second-order beliefs, when promises are made. Table 1 shows that this had the intended effect. Consider recipients who received and dictators who made a promise (columns (a) and (b), respectively). Their average first- and second-order beliefs are significantly higher when the chance of being re-matched is low rather than high: 0.74 is significantly higher than 0.55 ($Z = 2.20$, $p = 0.027$) and 0.80 is significantly higher than 0.66 ($Z = 2.20$, $p = 0.027$). Thus, a positive correlation between switching probability and beliefs exists when a promise was made.

This is not the case when no promise was made. Columns (c) and (d) show that first- and second-order beliefs are not affected by the switching probability.[14] Only if a promise were made did people take the value of the switching probability into account while forming expectations.

Table 1 also unveils a positive correlation between promises and expectations. On average, recipients who received a *Roll*-promise, and dictators who promised to choose *Roll*, have higher first- and second-order beliefs compared to those who did not receive/make such a promise. This positive correlation between promises and expectations confirms the findings of C&D and Vanberg. It exists both when the switching probability is high and when it is low.[15]

4. Main results

We now investigate the effects of expectations on dictators' promise-keeping behavior. According to H1, the causal factor leading to promise-keeping is a change in the expectations associated with a promise made. Then, we focus on 286 dictators who made a promise and were not switched (own promises). Columns of Table 2 report their average *Roll* rates (a), their second-order beliefs (c), and the first-order beliefs of recipients matched with them (b). The table also reports standard deviations and observations. Treatments with low switching probability (i.e. higher beliefs) are in row (1) while treatments with high switching probability (i.e. lower beliefs) are in row (2).

[13] Appendix A describes how beliefs in Table 1 (and in Tables B1 and B2 of Appendix B) were elicited and computed.

[14] In column (c), 0.42 is not significantly different from 0.41 ($Z = 0.73$, $p = 0.436$) and in column (d) 0.38 is not significantly different from 0.40 ($Z = 0.93$, $p = 0.345$).

[15] Regarding first-order beliefs, 0.74 is significantly higher than 0.42 ($Z = 2.20$, $p = 0.027$) and 0.55 is significantly higher than 0.41 ($Z = 2.20$, $p = 0.027$). Regarding the second-order beliefs, 0.80 is significantly higher than 0.38 ($Z = 2.20$, $p = 0.027$) and 0.66 is significantly higher than 0.40 ($Z = 2.20$, $p = 0.027$).

Table 3
Testing CBE & GA (575 obs.).

Switching probability	Recipients received		
	A promise (No Switch)	A promise (Switch)	
	Dictators sent a promise (a)	Dictators sent a promise (b)	Dictators sent no promise (c)
(1) Low (25%) [high SOBs]	51% (s.d. 0.50, obs. 214)	47% (s.d. 0.50, obs. 47)	29% (s.d. 0.46, obs. 21)
(2) High (75%) [low SOBs]	58% (s.d. 0.50, obs. 72)	28% (s.d. 0.45, obs. 173)	27% (s.d. 0.45, obs. 48)

The table reports average *Roll* rates of the dictators who face a recipient who received a promise from them (No Switch) or from others (Switch). Columns (b) and (c) split switched dictators based on what they did during the communication phase before being re-matched with a new recipient.

Test for EBE appears in column (a). Columns (b) and (c) report exogenous variation in beliefs in the subsample of non-switched pairs.[16] As in Table 1, both recipients' first-order beliefs and dictators' second-order beliefs are significantly higher when the chance of being re-matched is low rather than high.[17]

The *Roll* rates in column (a) are not higher when second-order beliefs are higher: the 51% *Roll* rate of those with high second-order beliefs is obviously not higher than the 58% *Roll* rate of dictators with low second-order beliefs ($Z = 0.93$, $p = 0.345$). Our data do not provide support in favor of EBE.

We now test CBE. According to H2, the casual factor leading to promise keeping is the act of making a promise. Thus, *Roll* rates should be higher when it comes to keeping own promises than promises made by others. The test for such effect is conducted along the *rows* of Table 3 where, holding switching probability constant, we can compare non-switched dictators who sent a promise with switched dictators facing a recipient who received a promise from someone else. Table 3 reports the average *Roll* rates pooling dictators in three groups: (*i*) non-switched dictators who sent a promise (column (a)); (*ii*) switched dictators who made a promise to their prior co-player and were re-matched with a new co-player who received a promise from someone else (column (b)); (*iii*) switched dictators who did not make a promise during their prior communication and were re-matched with a co-player who received a promise from someone else (column (c)).

Consistently with Tables 1 and 2, we can assume that dictator's second-order beliefs are the same in columns (a), (b), and (c) as the information sets of recipients are the same; i.e., their first-order beliefs are independent of the switch condition, while they only depend on the switching probability.

First, let us compare (a) and (b) in Table 3. Here any difference in the *Roll* rates can only be explained by the switch condition, because both switched and non-switched dictators made a promise in the first place. When the switching probability is high (second-order beliefs are low – row (2)), the fraction of non-switched dictators who chose to keep their promise is 58%, significantly larger than the 28% who chose to keep the promise of someone else in the switched group ($Z = 2.20$, $p = 0.027$). This test supports Vanberg's claim (2008, p. 1468) 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*." The difference in fact occurs despite dictators' average second-order beliefs not being significantly different between the two groups. By contrast, we find no evidence in favor of CBE when expectations are high (the switching probability is low – row (1)). Here the *Roll* rate of non-switched dictators is 51%, not significantly greater than the 47% of switched dictators ($Z = 0.52$, $p = 0.600$).

A comparison along the rows of columns (a) and (c) would not represent a clean test of CBE. Here the possible difference in behavior might be driven by the initial action of the dictator rather than the change in the switch condition.[18]

We can now test H3(a) along columns (b) and (c) of Table 3. Here we have switched dictators who sent (column (b)) or did not send (column (c)) a promise. Their re-matched recipients received a promise. The factor leading a change in second-order beliefs is a change in switching probability. Dictators' second-order beliefs are either high or low in rows (1) and (2), respectively. We can study how their behavior is affected by their beliefs. Consider column (b). Switched dictators sent a promise in the first place. Their average *Roll* rate is 47% when their second-order beliefs are high. It is significantly higher than the 28% *Roll* rate when their second-order beliefs are low ($Z = 1.99$, $p = 0.046$). This provides evidence is in favor H3(a). Now consider column (c). When switched dictators did not send a promise in the first place, we find no evidence in favor of GA (H3(a) is rejected). The average *Roll* rate of these dictators is 29% when second-order beliefs are high and 27% when second-order beliefs are low. They are not significantly different ($Z = 0.84$, $p = 0.399$).

We can further investigate GA of the switched dictators by using Table 4. We consider *switched* dictators who made a promise in the first place. We split them in two groups. Those who are re-matched with a recipient who received a promise in prior communication (column (a)), and those who are re-matched with a recipient who received no promise (column (b)). Here, by assumption, the exogenous variation in expectations is generated by prior communication. For any

16 Please notice the difference between Table 1 (col. (a) and (b)) and the data on first- and second-order beliefs in Table 2 (col. (b) and (c)). The former includes all pairs of subjects that exchanged promises during the communication phase – before they were eventually re-matched. In the latter, we only consider pairs in the no-switch condition.

17 Both recipients' first-order beliefs and dictators' second-order beliefs are significantly higher in row (1) than in row (2): 0.74 is significantly higher than 0.51 ($Z = 2.20$, $p = 0.027$) and 0.80 is significantly higher than 0.65 ($Z = 1.99$, $p = 0.046$).

18 Recall that in column (a) dictators initially sent a promise, while dictators in column (c) did not.

Table 4
Testing GA among switched dictators (289 obs.).

Switching probability	Recipients received	
	A promise [high SOBs]	No promise [low SOBs]
	Dictators sent a promise* (a)	Dictators sent a promise* (b)
(1) Low (25%)	47% (s.d. 0.50, obs. 47)	24% (s.d. 0.44, obs. 21)
(2) High (75%)	28% (s.d. 0.45, obs. 173)	27% (s.d. 0.45, obs. 48)

The table reports average *Roll* rates of switched dictators who sent a promise during the communication phase. Here recipients may or may not have been sent a promise during their prior communication with another subject before being re-matched (columns (a) and (b), respectively).

Table 5
Result summary.

Hypothesis	Table	Switching probability	Support
EBE (H1)	2	Low vs. high (no-switch)	No
CBE (H2)	3	Low (switch vs. no-switch)	No
		High (switch vs. no-switch)	Yes
GA (H3(a))	3	Low vs. high (switch)	Yes
GA (H3(b))	4	Low (promise vs. no-promise)	Yes
		High (promise vs. no-promise)	No

level of the switching probability, second-order beliefs are higher in the first group than in the second one. According to H3(b), holding switching probability constant, we compare average *Roll* rates in the two groups. We find evidence of GA along row (1). The 47% *Roll* rate when expectations are high is significantly higher than the 24% *Roll* rate when expectations are low ($Z = 1.99$, $p = 0.046$). We find no evidence of GA along row (2). The 28% of column (a) is not significantly different than the 27% of column (b) ($Z = 0.52$, $p = 0.600$).

Finally, in order to see whether the behavior of subject changes as they gain experience, we compared the first part of the experiment (rounds 1–4) with the second part (rounds 5–8). We found no substantial differences in the behavior of subjects. The only exception is CBE when switching probability is high. CBE does not hold in rounds 1–4, while it holds in rounds 5–8. Data are available from the authors upon request.

Table 5 summarizes our findings. The first column reports the hypotheses we test. The second column mentions the table where the test is reflected. The third column refers to the switching probability involved and the condition we consider to make comparisons of average *Roll* rates. The fourth column indicates whether or not the test supports the tested hypothesis.

5. Conclusion

Why do people keep their promises? The question tends on the brazen, obscuring how in many cases people actually renege. This is true in our data; in Table 2, the cell with the most frequent promise-keeping gets 58%, so 42% reneged. In other situations, reneging rates may even higher.[19] Yet, it is undeniably the case that in many contexts, including the one we studied, promises can have a dramatic effect on trust and cooperation, even if not foolproof. It is important to understand why.

The expectations-based explanation (EBE) of promise-keeping says that people have a general tendency to fulfill others' expectations, and that they will keep a promise if it substantially increased their counterpart's expectations. Commitment-based explanation (CBE) says that people attach a value to the act of keeping their word, *per se*. Thus, people will exhibit a higher inclination to keep their promises than the promises made by others. We argued that in principle the two explanations are not mutually exclusive, since the inclination to keep own promises may still positively depend on expectations.

We consider the prominent lab-context introduced by Vanberg (2008), with rich free-form communication followed by a potential "partner-switch." We make the probability of a partner-switch a treatment variable, thereby inducing high vs. low recipient expectations (for low vs. high switch-probability). Differently from previous work, our design allows us to test the empirical relevance of both EBE and CBE.

Our finding supports the notion that the value a person attaches to keeping her promise is independent of her beliefs. When it comes to keeping own promises, our experiment offers no support in favor of EBE. Our data on switched dictators,

[19] Charness and Dufwenberg (2010) suggest that, outside the lab, this includes "used car sales, promises made by politicians, tax returns sent to the IRS, and testimony in traffic courts." And in-the-lab they show that this is the case when promises are "bare" (viz., circling a pre-fab message) as opposed to "rich" (free-form, as in C&D). Realistically, intrinsic motivations to keep one's word or to live up to others' expectations are not unconditional. Not only do motivations incorporate absolute principles of justice, they also depend on contextual factors which may strengthen or weaken them (e.g., habits, norms, culture, merit, need, ...). How these factors influence the pro-social choice of keeping a promise is an important issue yet to be unveiled by future research.

however, suggests that beliefs may still affect the behavior of those who have to choose whether to keep a promise made by someone else.

Appendix A. Elicitation of beliefs

Elicitation of first-order beliefs: After communication, recipients were asked to guess what their (unknown) dictators would choose to do. They had been told the value of the switching probability in their treatment. Thus, they were aware that their paired subject could be switched according to that probability. Recipients could make their guess by ticking one of the five-point scale in Table A. This scale is the same as in Vanberg. Beliefs are then re-scaled to 1, 0.75, 0.5, 0.25, and 0. Thus the numbers shown in Table 1 (and in Tables B1 and B2 below) represent the averages of recipients' re-scaled responses. The payoffs correspond to a quadratic scoring rule for probability values 85%, 68%, 50%, 32%, and 15%, because due to the risk neutrality assumption, quadratic scoring yields flat payoffs as probabilities approach one (see Vanberg, p. 1472).[20]

Table A
Incentives for first-order belief elicitation.

	The dictator will				
	choose *Roll*			choose *Don't Roll*	
	Certainly	Probably	Unsure	Probably	Certainly
Please tick your guess	◯	◯	◯	◯	◯
Your earnings if the dictator					
chooses *Roll*	0.65 tokens	0.60 tokens	0.50 tokens	0.35 tokens	0.15 tokens
chooses *Don't Roll*	0.15 tokens	0.35 tokens	0.50 tokens	0.60 tokens	0.65 tokens

Elicitation of second-order beliefs: Before dictators were told whether their paired subject had been switched or not, they were asked to guess his guess. Specifically, they had to guess which of the five points of Table A had been ticked by their counterpart. Correct guesses were paid 0.50 tokens.

Appendix B. Beliefs

As shown in Table B1, recipients' first-order beliefs are independent of whether pairs were re-matched or not. This result is of course consistent with the fact that recipients knew that only dictators would have been informed of a possible switch, in the case it occurred. Consider the first row. Beliefs of recipients who received a promise are 0.74 in the no-switch condition and 0.75 in the switch condition ($Z = 0.52$, $p = 0.600$). Beliefs of recipients who received no promise are 0.41 in the no-switch condition and 0.44 in the switch one ($Z = 0.10$, $p = 0.916$). Similar results hold in the second row. Beliefs of recipients who received a promise are 0.51 in the no-switch condition and 0.57 in the switch one ($Z = 0.94$, $p = 0.345$). Beliefs of those who received no promise are 0.41 in the no-switch condition and 0.41 in the switch one ($Z = 0.10$, $p = 0.916$).

Table B1
First-order beliefs (obs. 768).

Switching probability	Received a promise		Did not receive a promise	
	No switch (1)	Switch (2)	No switch (3)	Switch (4)
Low (25%)	0.74 (s.d. 0.30, obs. 214)	0.75 (s.d. 0.31, obs. 68)	0.41 (s.d. 0.37, obs. 74)	0.44 (s.d. 0.42, obs. 28)
High (75%)	0.51 (s.d. 0.36, obs. 72)	0.57 (s.d. 0.34, obs. 221)	0.41 (s.d. 0.34, obs. 24)	0.41 (s.d. 0.31, obs. 67)

The data reported in the columns of Table B1 is also consistent with the exogenous variation in expectations described in Table 1:

1. First-order beliefs of recipients who received no promise appear to be independent of the switch condition. In the no-switch condition they are 0.41 when the switching probability is low and 0.41 when it is high ($Z = 0.11$, $p = 0.916$). First-order beliefs in the switch condition are 0.44 when the switching probability is low and 0.41 when it is high ($Z = 0.11$, $p = 0.916$).
2. By contrast, average first-order beliefs of recipients who received a promise depend on the switching probability. In the no-switch condition they are 0.74 when the switching probability is low and 0.51 when it is high ($Z = 2.20$, $p = 0.027$). In the switch condition they are 0.75 when the switching probability is low and 0.57 when it is high

[20] We also verify the robustness or our results to the quadratic scoring rule. Results are available upon request.

Table B2
Second-order beliefs (obs. 768)*.

Switching probability	No switch (384 obs.)		Switch (384 obs.)	
	Promised to Roll (1)	Did not promise (2)	Promised to Roll (3)	Did not promise (4)
Low (25%)	0.80 (s.d. 0.27, obs. 214)	0.38 (s.d. 0.38, obs. 74)	0.79 (s.d. 0.26, obs. 68)	0.38 (s.d. 0.38, obs. 28)
High (75%)	0.65 (s.d. 0.35, obs. 72)	0.47 (s.d. 0.35, obs. 24)	0.67 (s.d. 0.33, obs. 221)	0.37 (s.d. 0.35, obs. 67)

* Note that second-order beliefs were elicited before dictators knew whether they had been switched or not. The values in the table have been split in *no-switch* and *switch* by the experimenter.

($Z = 2.20$, $p = 0.027$). As expected, these beliefs are independent of whether pairs were re-matched. Moreover, all these values are significantly higher compared to those of recipients who did not receive a promise.

Table B2 presents dictators' average second-order beliefs.

When the chance of being switched is low, dictators who promised to *Roll* (columns (1) and (3)) display higher average second-order beliefs: 0.80 is significantly higher than 0.65 ($Z = 1.99$, $p = 0.046$) and 0.79 is significantly higher than 0.67 ($Z = 2.20$, $p = 0.027$).

The value of the switching probability does not affect second-order beliefs of dictators who are matched with recipients that received no promise (cf. columns (2) and (4)): 0.38 is not significantly different from 0.47 ($Z = 0.94$, $p = 0.345$) and 0.38 is not significantly different from 0.37 ($Z = 0.52$, $p = 0.600$). This is consistent with the fact that recipients who received no promise do not care about who they will be paired with. Since no promise has been made, they expect dictators are equally likely to choose *Don't Roll*, independently of whether they have been re-matched or not, and independently of the switching probability.

Reading Table B2 by "rows" indicates that promises are positively correlated with expectations. Such correlation exists both when the switching probability is high and when it is low.[21]

Remember that the average second-order beliefs reported in Table B2 refer to the original partner. Here the switching condition does not matter since second-order beliefs are elicited before the switch occurs. They are high when the dictator makes a promise and they depend on the switching probability as expected for the exogenous variation in expectations. These are the relevant second-order beliefs when we look at the behavior of non-switched dictators. By contrast, they are not relevant when we consider the behavior of the switched dictators. In this case, as we explain in the design section, the relevant second-order beliefs are those referring to their new partners. Since we do not elicit them, we assume that the exogenous variation in beliefs occurs also with reference to the expectations of the new partners.

Appendix C. No promise

What happens when no promise was made during the communication phase? As expected, there is strong correlation between the existence of a promise and *Roll* rates, a result in line with C&D and Vanberg. Table C summarizes average *Roll* rates when no promise was made. It distinguishes between switch and no-switch conditions. It also reports standard deviations (s.d.) and observations (obs.). *Roll* rates of no switch dictators drop substantially compared to when a promise was made (cf. Table 2). The table also shows a significant difference between the average behavior of dictators who made no promise and were not switched (0.27) and dictators who were re-matched with a new recipient who had received no promise (0.15; $Z = 1.99$, $p = 0.046$).

Table C
Dictators' behavior (*Roll* rates). Partners received no promise (obs. 124).

Second-order beliefs	No switch (a)	Switch (b)
Lowest SOBs*	0.27 (s.d. 0.44, obs. 98)	0.15 (s.d. 0.37, obs. 26)

* When no promise was made we obtain the lowest SOBs (see Appendix B, Table B2 – columns (2) and (4)). Appendix B also shows that the change in the switching probability does not yield any significant variation in expectations.

Appendix D. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.geb.2018.07.009.

[21] Among non-switched dictators: 0.80 vs. 0.38 ($Z = 2.20$, $p = 0.027$) and 0.65 vs. 0.47 ($Z = 1.99$, $p = 0.046$), in the low and the high switching probability case, respectively. Among switched dictators: 0.79 vs. 0.38 ($Z = 2.20$, $p = 0.027$) and 0.67 vs. 0.37 ($Z = 2.20$, $p = 0.027$), in the low and the high switching probability case, respectively.

References

Bacharach, M., Guerra, G., Zizzo, D.J., 2007. The self-fulfilling property of trust: an experimental study. Theory Dec. 63, 349–388.
Battigalli, P., Dufwenberg, M., 2007. Guilt in games. Amer. Econ. Rev. 7, 170–176.
Bellemare, C., Sebald, A., Strobel, M., 2011. Measuring the willingness to pay to avoid guilt: estimation using equilibrium and stated belief models. J. Appl. Econometrics 26 (3), 437–453.
Bicchieri, C., 2006. The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press, New York.
Bracht, J., Regner, T., 2013. Moral emotions and partnership. J. Econ. Psych. 39, 313–326.
Braver, S., 1995. Social contracts and the provision of public goods. In: Schroeder, D. (Ed.), Social Dilemmas: Perspectives on Individuals and Groups. Praeger, New York.
Chang, L.J., Smith, A., Dufwenberg, M., Sanfey, A., 2011. Triangulating the neural, psychological and economic bases of guilt aversion. Neuron 70, 560–572.
Charness, G., Dufwenberg, M., 2006. Promises and partnership. Econometrica 74, 1579–1601.
Charness, G., Dufwenberg, M., 2010. Bare promises: an experiment. Econ. Letters 107, 281–283.
Dufwenberg, M., Gneezy, U., 2000. Measuring beliefs in an experimental lost wallet game. Games Econ. Behav. 30, 163–182.
Ederer, F., Stremitzer, A., 2017. Promises and expectations. Games Econ. Behav. 106, 161–178.
Ellingsen, T., Johannesson, M., 2004. Promises, threats and fairness. Econ. J. 114, 397–420.
Ellingsen, T., Johannesson, M., Tjøtta, S., Torsvik, G., 2010. Testing guilt aversion. Games Econ. Behav. 68, 95–107.
Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready–made economic experiments. Exper. Econ. 10, 171–178.
Guerra, G., Zizzo, D.J., 2004. Trust responsiveness and beliefs. J. Econ. Behav. Organ. 55, 25–30.
Khalmetski, K., Ockenfels, A., Werner, P., 2015. Surprising gifts: theory and laboratory evidence. J. Econ. Theory 159, 163–208.
Krupka, E.L., Leider, S., Jiang, M., 2017. A meeting of the minds: contracts and social norms. Manage. Sci. 63, 1708–1729.
Ostrom, E., Walker, J., Gardner, R., 1992. Covenants with and without a sword: self-governance is possible. Amer. Polit. Sci. Rev. 86, 404–417.
Regner, T., Harth, N.S., 2014. Testing Belief-Dependent Models. WP, Max Planck Institute of Economics, Jena.
Reuben, E., Sapienza, P., Zingales, L., 2009. Is mistrust self-fulfilling? Econ. Letters 104, 89–91.
Vanberg, C., 2008. Why do people keep their promises? An experimental test of two explanations. Econometrica 76, 1467–1480.