Promises or Agreements?
Moral commitments in bilateral communication

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

October 20, 2022

Abstract. In a game with pre-play bilateral communication, messages may trigger moral incentives to honor promises or agreements. We hypothesize that individuals' inclination to keep a promise is highest if the counterpart requited the promise. We interpret this as an inclination to honor agreements. We report supporting results from an experiment.

Keywords: Guilt aversion, promise-keeping, informal agreements.

JEL codes: A13, C91, D03, D64, D90.

[a] Department of Economics and Law, Sapienza University of Rome, Italy; Department of Economics, University of Antwerp, Belgium.
[b] Department of Economics, University of Arizona, US; Department of Economics, University of Gothenburg, Sweden; CESifo, Munich, Germany.
[c] Department of Economics and Finance, University of Tor Vergata, Rome, Italy.
[d] Department of Economics - Esomas, University of Turin, Italy; Collegio Carlo Alberto, Italy; Bocconi University, Milan, Italy; CESifo, Munich, Germany.

1. Introduction

Vanberg (2008) shows that people's taste for keeping their word can be an important driver of pro-social behavior. His context was one where subjects exchange messages. Hence, a player's promise may or may not be requited, although Vanberg did not focus on that. We explore an experimental design which is similar to Vanberg's and hypothesize that individuals' inclination to keep promises is highest when their promises are requited. We interpret such a preference as an inclination to honor agreements.

2. Hypotheses

Vanberg's game is presented in Figure 1. Players 1 and 2 have an equal chance to be assigned the role of dictator, i.e., the player who chooses between *Don't Roll* and *Roll*. The given payoffs reflect monetary payments, not necessarily utilities, as individual choices may be affected by social preferences (e.g., inequity aversion, reciprocity, pangs of guilt, …).
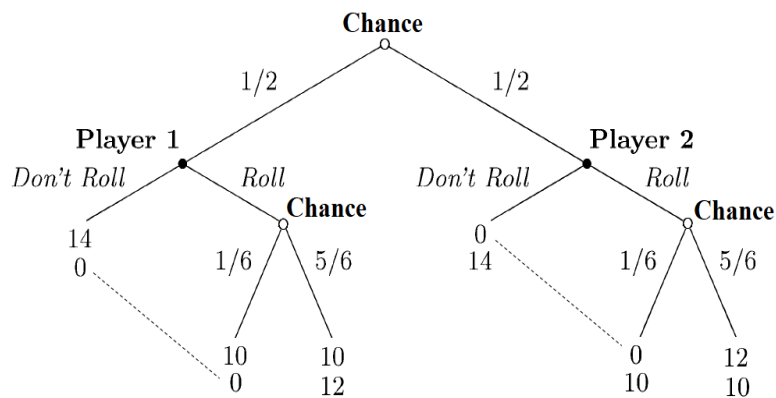

Figure 1 – Vanberg's game

Before the players are told their roles, they can exchange up to four messages. One player is randomly chosen to send the first message. After communication, players are randomly assigned to their roles, dictator or recipient. Furthermore, *half of the recipients are re-matched with a new dictator*. Only dictators observe the switch; moreover, switched dictators can read the messages between their new recipient and his/her counterpart before re-matching.[1] Finally, each dictator chooses between *Roll* and *Don't Roll*, as shown in Figure 1.

---

[1] These features imply that dictators' second-order beliefs are switch-independent, which is essential for ruling out confounds created if subjects are guilt-averse as considered by Charness and Dufwenberg (2006) (and compare also Battigalli & Dufwenberg 2007). We do not focus on guilt aversion here. Refer to Vanberg (2006) and Di Bartolomeo *et al*. (2019) for more discussion.

Let us formulate a simple theory. Assume that players' utilities are affected by non-monetary motivations, which may depend on communication history. Namely, let the dictator's utility when he chooses *Don't Roll* be $14-m^d$, where m^d is the psychological/moral cost of not being generous. Of course, the dictator will choose *Roll* if $m^d>4$. Vanberg (and others) have reported evidence supporting that m^d is higher when a promise has been made. We propose that m^d is higher when the promise has been made as a part of an agreement than an unrequited promise.

We test this idea in two ways. First, we focus solely on dictators whose recipient was not switched. We compare the *Roll* rates of those who reached an agreement (RR-A-NS for *Roll*-rate-agreement-no-switch) to with that of those whose promise was not requited (RR-UP-NS for *Roll*-rate-unrequited-promise-no-switch). Apply the above theory and allow for some individual variation as regards levels of m^d; we arrive at the following hypothesis:

H1:   RR-A-NS > RR-UP-NS

Support for H1 would be consistent with our idea that agreements create more moral commitment than unrequited promises. However, such a pattern could also be consistent with beliefs-driven motivations (compare footnote 1). Vanberg introduced his switching-feature to cater to a similar confound. We follow his lead and arrive at our second test. We compare agreement-forming dictators depending on whether they were subsequently switched. Specifically, compare RR-A-NS (defined as above) to the *Roll* rate of those dictators who initially reached an agreement but were switched and then matched with a new recipient who had previously formed an agreement with someone else (RR-A&A-S).[2] The idea, analogous to Vanberg's but here applied to agreements as opposed to any promise, is that m^d is higher for the former group of dictators than for the latter. On choosing *Don't Roll,* the latter group would not view themselves as violating an agreement, since the person with whom they had an agreement was switched. Proceeding as with H1, we arrive at the following hypothesis:

H2:   RR-A-NS > RR-A&A-S

---

[2] The focus on new recipients that previously formed an agreement (as opposed to any new recipients) is crucial to ensure that dictators' second-order beliefs are switch-independent. Compare footnote 1.

3. Experimental design[3]

The design involved 192 undergraduate students. Each played a session. We performed 6 sessions. Each session involved 32 participants and consisted of 8 rounds with perfect stranger matching.

In each round, participants played the game described in Figure 1. Before playing and knowing their roles, subjects chatted.

Each round implemented the following sequence stages.

1. **Communication.** Subjects were randomly matched to form 16 chatting pairs. Each chat consisted of four one-way max-90-character messages in sequence with a random determination of who would start to chat.

2. **Role assignment.** After communication, roles were randomly assigned in each pair.

3. **Switching.** 50% of recipients were re-matched with new dictators. Only dictators were informed whether a switch occurred. Switched dictators were allowed to read the prior conversation of their new recipients.

4. **Dictators' action.** Dictators made their choice: *Roll* or *Don't Roll*. After, participants were informed of their round payoffs. Recipients were not informed whether they had been switched, nor could they infer the dictator's choice when their payoffs were zero.[4]

At the end of each session, one of the rounds was randomly chosen for payments determined by dictators' choices. All the payoffs were described in tokens, with 1 token = 0.5 euros. Each participant also received a fixed show-up fee of 2.50 tokens.

4. Experimental results

Our sample consists of 768 chats.[5] Out of those, we obtained 395 agreements (51%), where both parties promised to *Roll*; 204 cases where promises were not requited (27%); 169 chats (22%) ended with no promise.

---

[3] For the sake of brevity, here, the experimental design is just sketched. Instructions and additional tests are available upon request.

[4] 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".

[5] As said, following Vanberg (2008), we consider each player's chat a single message. Research assistants, blind to our hypotheses, catalogued all the messages. The research assistant was not involved in the design and execution of the experiment. Indeed, we asked three research assistants to classify messages and *ex ante* randomly choose the classification of one of them for the experiment. The different classifications were, however, strongly correlated (with a high Cronbach-alpha value of 0.8604.)

Our results are described in Table 2. It reports the *Roll* rates for switched and non-switched dictators. Rows give information about the kind of communication that occurred. Standard deviations/number of observations are reported in brackets.

Table 2 – No-switched and switched SOBs and *Roll* rates

Communication outcomes	ROLL RATES	
	(a) NO SWITCH	(b) SWITCH[*]
(1) AGREEMENT	0.59	0.40
	(0.49/198)	(0.49/144)
(2) DICTATOR'S PROMISE NOT REQUITED	0.43	0.47
	(0.50/42)	(0.52/43)
(3) NO PROMISE	0.25	0.27
	(0.43/97)	(0.45/37)
(4) RECIPIENT'S PROMISE NOT REQUITED	0.34	0.24
	(0.48/47)	(0.44/25)

(*) As in Vanberg (2008), to avoid potential self-selection bias, we only consider switched dictators who promised to *Roll* in rows (1)-(2), while we only consider switched dictators who did not make any promise in rows (3)-(4). In each row, SOBs in (a) and (b) are not statistically different. As a result, the table displays the outcomes from 633 out of 768.

Let us begin with H1. We find that non-switched dictators who formed an agreement *Roll* more frequently (59%) than those whose promise was not requited (43%): $Z=1.99$, $p=0.046$.[6] Hence, our outcomes are consistent with our idea that agreements create more moral commitment than unrequited promises.

The support for H1, however, could also be driven by subjects' guilt-aversion. In H2 we then rule out these confounds. Non-switched dictators who achieved an agreement are more likely to *Roll* (59%) than those dictators who initially reached an agreement but were switched and then matched with a new recipient who had previously formed an agreement with someone else (40%): $Z=2.20$, $p=0.028$. As of H2, Vanberg's moral commitment to keep one's word seems to bind for Agreements.

We can look at the other cells of our table. Non-switched dictators whose promise was not requited *Roll* more (43%) than those who did not make any promise (25%). However, although the difference in probability is large, the effects of non-responded promises are statistically not different from zero. Hence, promises not requited are not statistically different from empty communication.

---

[6] All the statistics are obtained using the Wilcoxon signed-rank test, which compares averages at the session level. Our data are independent at the session level, but not at the individual level.

References

Battigalli, P. & M. Dufwenberg (2007), "Guilt in games," *American Economic Review,* 97: 170–176.

Charness, G. & M. Dufwenberg (2006), "Promises and partnership," *Econometrica*, 74: 1579–1601.

Di Bartolomeo, G., M. Dufwenberg, S. Papa & F. Passarelli (2019), "Promises, expectations & causation," *Games & Economic Behavior*, 113: 137–146.

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