Journal of the Economic Science Association (2024) 10:578–594
https://doi.org/10.1007/s40881-024-00180-8

ORIGINAL PAPER

# Guilt aversion and moral commitment: Eve versus Adam

**Giovanni Di Bartolomeo**[1,5,6] · **Martin Dufwenberg**[2,7,8] · **Stefano Papa**[3,6] · **Laura Razzolini**[4,9]

Received: 18 October 2023 / Revised: 22 February 2024 / Accepted: 21 August 2024 /
Published online: 3 October 2024


## Abstract
We explore gender differences in individuals' motivations. We focus on guilt aversion and moral commitment. Our experiment supports the idea that men are more guilt-averse than women, while moral motivations drive more women's actions in a random dictator game with pre-play communication.

**Keywords** Gender · Guilt aversion · Promises · Evolutionary psychology

**JEL Classification** A13 · C91 · D03 · D64

✉ Giovanni Di Bartolomeo
giovanni.dibartolomeo@uniroma1.it

1 Department of Economics and Law, Sapienza University of Rome, Roma, Italy

2 Department of Economics, University of Arizona, Tucson, US

3 Department of Economics and Finance, University of Rome Tor Vergata, Roma, Italy

4 Department of Economics, Finance and Legal Studies, University of Alabama, Tuscaloosa, US

5 Department of Economics, University of Antwerp, Antwerp, Belgium

6 CIMEO, Sapienza University of Rome, Rome, Italy

7 University of Gothenburg, Gothenburg, Sweden

8 CESifo, Munich, Germany

9 Center for the Philosophy of Freedom, University of Arizona, Tucson, US

## 1 Introduction

Behavioral and experimental economists have defined and tested guilt aversion (GA) and moral commitment (MC) in games, trying to understand and explain trustworthy behavior in the[1]. This literature, however, has not extensively addressed potential gender differences in GA or MC. We experimentally explore whether GA and MC affect men differently than women in the laboratory. As the directional nature of such an effect is not self-evident from the few empirical studies available,[2] we do not propose directional hypotheses for our tests on gender differences.

Inspired by Vanberg (2008), we use a simple dictator game with pre-play communication to disentangle the non-monetary (potentially different) motivations that drive women's and men's actions. We distinguish motivations driven by belief-dependent GA from motivations driven by MC, which is independent of beliefs. Testing for differences in motivations is a complex task for two reasons: i) GA- and MC-driven motivations are not mutually exclusive; ii) different motivations imply different causations even though both are compatible with the same correlation between actions and expectations.

We propose a set of tests based on exogenous variations and difference-in-difference comparisons to address the above-described challenging task. In what follows, we sketch our approach and attempt to provide some intuition; more details will be introduced in later sections.

In our experiment, pairs communicate with each other and exchange promises about their potentially generous behavior in a dictator game. After communicating, the pair plays a dictator game, with one individual chosen to be the dictator and the other to be the recipient. However, with a given known probability (revealed after the communications stage is concluded), some pairs are switched, so a dictator can play with a recipient different from the one with whom he/she had previously communicated. Furthermore, the switch is revealed only to the dictators. Hence, the recipients' expectations, their first-order beliefs (FOBs), and the dictators' second-order beliefs (SOBs) about the recipients' beliefs are independent of the switch.

Promises fuel expectations independently of the driving motivations. Therefore, on average, recipients who experienced promises would have higher expectations when the switching probability is low than those who experienced promises in a high switching probability scenario. In this way, we create an exogenous variation in beliefs with high or low SOBs, as dictators know that recipients do not observe whether a switch occurred. Different implications for GA can, then, be searched in the data.

Similarly, we can glean variation in intrinsic motivations by comparing, everything else equal, the behavior of a dictator asked to keep their own promise to

[1] Battigalli and Dufwenberg (2007) develop a general theory. Dufwenberg and Gneezy (2000) and Charness and Dufwenberg (2006) ran early experiments. See, among others, Khalmetski et al. (2015), Bellemare et al. (2017), Ederer and Stremitzer (2017), Attanasi et al. (2019), Dhami et al. (2019), Di Bartolomeo et al., (2019a, 2019b, 2023a, 2023b), Ciccarone et al. (2023), Charness et al. (2023), Cartwright (2019) for subsequent experimental literature. See Battigalli and Dufwenberg (2022, especially Sects. 3.1 and 7) for a broad related discussion.

[2] The existing literature is discussed later in the introduction.

the behavior of a dictator asked to keep the promise made by someone else. Here, MC implies that we are more likely to observe promise-keeping by the former dictators than the latter. We separately explore MC and GA by gender and test how each affects behavior across genders using a difference-in-difference approach.

In the case of MC, we focus on its marginal effect, i.e., the difference in the behavior of promisors asked to keep their own promises to that of promisors asked to keep a promise made by someone else. Then, we compare the marginal effects for men and women, respectively. Similarly, the marginal effect of GA is measured by the change in behavior due to a change in SOBs, and then, we test by gender for differences in marginal effects.

Let us compare our approach to some existing work. Vanberg (2008), Di Bartolomeo et al. (2019b), and Kleinknecht (2019) are methodologically closest to ours as they share the same experimental setup: a random dictator game with pre-play communication:

Vanberg (2008) conducted a pioneering study using an exogenous variation technique to test for the effects of MC and GA. He used a fixed 0.5 switching probability to compare the behavior of dictators who were asked to keep their own promises to the behavior of those asked to keep promises made by others. Vanberg (2008) found support in favor of MC. Vanberg (2008) also investigated how the behavior of dictators who made promises varied when their recipients had or had not received a promise from someone else. He looked for evidence in support of GA given that dictators playing with promise-receiving recipients held higher subjective second-order beliefs (SOBs). Vanberg did not find support for GA. Kleinknecht (2019) replicated Vanberg's (2008) experiment and did find support for GA. She also controlled for gender and found no significant differences in MC between men and women.

Di Bartolomeo et al. (2019b) expanded upon Vanberg (2008) and Kleinknecht (2019) using different switching probabilities (high and low) to create a double exogenous variation. They tested for both GA and MC-driven motivations and found evidence in support of both. Additionally, Vanberg's and Kleinknecht's approaches have the questionable feature that they compared dictators based on whether their new recipients received a promise. This approach involves two variables: the SOBs of dictators and the promise status of recipients. This double comparison is not ideal from an experimental point of view. Instead, Di Bartolomeo et al. (2019b) addressed the issue by comparing dictators who played with recipients who differ only in their expectations and not in their promise-receiving status.

Compared to the above studies, our paper provides a double value added. First, we extend Kleinknecht (2019) by considering a double exogenous variation in SOBs and promises. As a result, we cleanly isolate and test for both MC- and GA-driven motivations within and between genders. Second, we explore gender sensitivity to GA (or MC) by considering a difference in difference.

Our study also contributes to the extensive literature in experimental economics that explores gender differences in how individuals are motivated.[3] Specifically, it is close to two papers which explore gender differences in GA with opposing results. Nihonsugi et al. (2022) recently presented tests designed to study gender differences in GA using psychological survey questions and a trust game. They found that men are more guilt-averse than women. Our results are in line with those of Nihonsugi et al. (2022) although they use a different experimental design and do not allow for communication. By contrast, Else-Quest et al. (2012) report some evidence in the opposite direction. They found that women are more guilt prone than men.[4]

Finally, our paper is related to the rich literature on GA and MC focused on eliciting GA parameters at the individual level. While these studies do not primarily focus on gender differences, they do report some findings, and most of the measured differences are insignificant (see Bellemare et al., 2018; Bellemare & Sebald, 2023).[5] The non-belief-dependent motivations relevant to our point of view (about MC) are discussed and surveyed, among others, by Ellingsen and Johannesson (2004) and Vanberg (2008). Promises are thought to create commitments to fulfilling contractual obligations.

Sections 2 and 3 present our experimental design, hypotheses, and procedures. Section 4 shows our results. Section 5 offers a concluding discussion.

## 2 The experiment and our hypotheses

### 2.1 The game

We use a two-player dictator game augmented with a) bilateral pre-play communication, b) random role assignment and asymmetric information, and c) a partner-switching mechanism. It works in five steps:

1) Participants are matched in pairs. Each pair can communicate by sending messages. The subjects know they will play a random dictator game, but will know later who will act in which role.
2) After the communication phase, each player is assigned the role of dictator or recipient.
3) After the communication phase and before making the allocation decision, a given proportion of dictators have their partners switched. Both players know the

[3] For pioneering efforts, see Bolton and Katok's (1995) and Eckel and Grossman's (1998) studies of generosity. Croson and Gneezy (2009) surveyed many papers that explored risk preferences, social preferences, or attitudes to competition. Other examples include Dreber and Johannesson (2008) and Childs (2012), who report results regarding attitudes to lying.

[4] See also Crowley (1997), Benetti-McQuoid and Bursik (2005), and McQuoid and Bursik (2005). These studies mainly focus on individuals' reactions from experiencing guilt rather than on how strongly the sentiment of GA is experienced.

[5] Insignificant findings, of course, do not necessarily mean that these papers show an absence of gender effects.

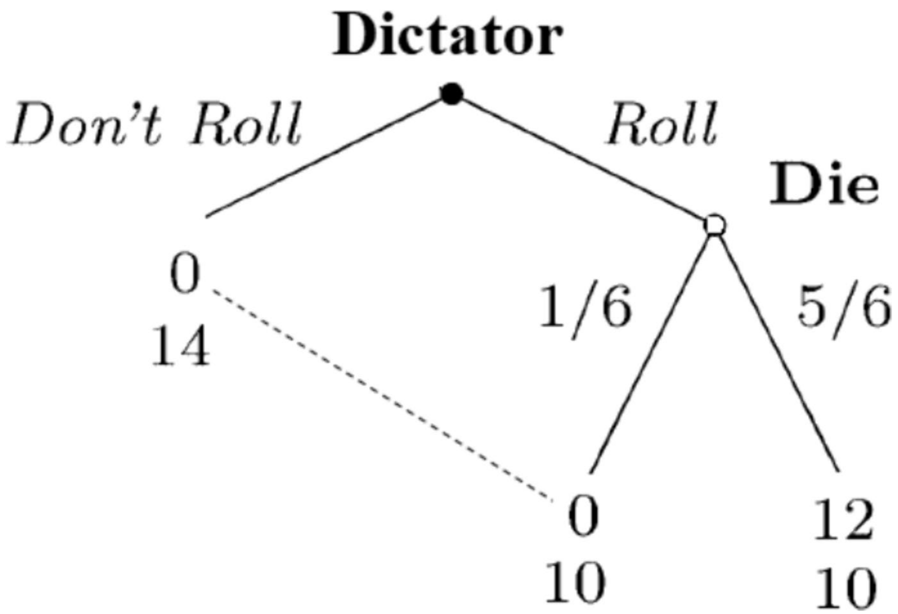


**Fig. 1** Game payoffs

proportion of switched pairs, but only dictators are told whether their partner has been switched. The recipients are not informed whether a switch occurred.

4) After the switch, dictators can read the messages (sent by another dictator) received by their new partners.
5) Each pair plays the game (form) shown in Fig. 1. Dictators choose *Roll* or *Don't Roll*. The payoffs are in euros, and the recipient's payoff is listed on top. The information set "00" indicates that the recipient is not informed how the payoff of "0" realized.

## 2.2 Hypotheses

Our exogenous-variation-based tests involve a subgroup of dictators. Specifically, we focus on dictators who made and "read" promises and refer to this group as Γ-dictators. Non-switched Γ-dictators must decide whether to keep their own promises, while switched Γ-dictators are promisors who are re-matched with a recipient who received a promise from another dictator. As recipients are not informed whether a switch occurred, FOBs and SOBs are independent of the switch. However, these beliefs can (and we show that they do) depend on the switching probability: a low (high) switching probability is associated with high (low) beliefs. Unless we explicitly say otherwise, we always refer to Γ-dictators. Thus, our paper only speaks about GA or MC of those subjects who have made a promise.

### 2.2.1 Exogenous variations

Our design leads to different exogenous variations, which can be used to test for GA (as in Di Bartolomeo et al., 2019b) and MC (as in Vanberg, 2008). We begin by introducing our hypotheses about SOBs. We denote by $S_i(j)$ the average SOBs of a dictator who is $i \in \{M[ale], F[emale]\}$, in a game with $j \in \{L[ow], H\left[igh\right]\}$ switching probability.

We start with three preliminary hypotheses that are crucial to all that follow since the subsequent hypotheses are conditional on their relevance. We explore whether there is any difference in the SOBs between women and men for any given switching probability (H1a) and whether our exogenous variations in SOBs work well for both genders (H1b). Finally, we wonder whether the potential variation of expectations observed for men and women is of the same magnitude for both genders, i.e., whether the variation in the probability of being switched (from high to low) generates the same change in expectations among individuals involved in a promise (H1c).

Formally:

**H1a (no difference in expectations between genders):** Subjects' expectations are independent of gender: $S_M(j) = S_F(j)$ for $j \in \{L, H\}$.

**H1b (exogenous variations in beliefs within genders):** Expectations of subjects involved in a promise are higher when the switching probability is low, for both men and women: $S_i(L) > S_i(H)$ for $i \in \{M, F\}$.

**H1c (marginal exogenous variations in beliefs between genders):** Changes in the switching probabilities imply the same changes in expectations for men and women initially involved in a promise: $S_F(L) - S_F(H) = S_M(L) - S_M(H)$.

In the case of H1a and H1c, we have no directional expectation. By contrast, H1b is expected to hold based on the extensive literature supporting a correlation between promise-keeping behavior and SOBs (for related results and discussion, see Vanberg, 2008; Di Bartolomeo et al., 2019b).

It should be noted that we elicit expectations before the switch occurs, so our tests are based on the expectations (FOBs and SOBs) regarding the actual communications that took place. We assume that the same beliefs would have been observed when subjects are asked to evaluate communication between others, meaning that a dictator involved in communication would evaluate it similarly even if they were not directly involved.[6] We elicited expectations beforehand to avoid distortions in elicitation due to timing and potential confusion generated in subjects who read their communications first and then the communication made by others (when switched) and were asked to evaluate the latter. This assumption only pertains to switched dictators, who are assumed to have—everything else equal—the same beliefs as the non-switched ones. Note that this assumption is supported by similar studies using a different procedure (after-communication elicitation, see e.g., Vanberg, 2008).

[6] Potentially, subjects could suffer from some form of self-confirmation bias. However, Vanberg (2008) finds that this is not the case.

### 2.2.2 Guilt aversion and gender

We focus on switched subjects to explore GA. In this way, we can abstract from any MC motivations for subjects' behaviors and concentrate on GA only. To formalize our hypotheses, we define by $R_i(j,k)$ the average *Roll* rate of $i \in \{M[ale], F[emale]\}$ Γ-dictator, who played a game with a $j \in \{L[ow], H\left[igh\right]\}$ switching probability, who has been $k \in \{S[witched], N[otswitched]\}$.

We test two main hypotheses. The first one is a clear implication of GA. Assuming that H2 holds, since a Γ-dictator's SOB is higher when the switching probability is low, if said г-dictator is guilt-averse, he/she is more likely to *Roll*.[7] Formally:

**H2 (implication of GA within genders):** Switched г-dictators are more likely to *Roll* when the switching probability is low: $R_i(L,S) > R_i(H,S)$ for $i \in \{M,F\}$.

We shall test H2 for each gender.

Furthermore, we explore differences in GA sensitivity between genders. We compare the potential increase in the average *Roll* rate for men associated with a rise in SOBs to the equivalent potential increases for women. Formally, assuming H1b and H1c hold, we use difference-in-difference in GA sensitivity between men and women as follows:

**H3 (guilt sensitivity):** In the case of low switching probability (high SOBs), male and female switched Γ-dictators are equally likely to *Roll*: $R_M(L,S) - R_M(H,S) = R_F(L,S) - R_F(H,S)$.

Note that the above test makes sense if the exogenous variation increased SOBs independent of gender (see H1c.) Note also that H2 is a directional hypothesis, while H3 is not. We will, therefore, use one-tailed tests for the former and two-tailed tests for the latter. As argued in the introduction, we do not impose any directional hypothesis for H3, as different theories may suggest opposite conclusions.

### 2.2.3 Moral commitment and gender

We need to reintroduce non-switched dictators to test MC-driven motivations. As in Vanberg (2008) and Kleinknecht (2019), we will compare the behavior of dictators asked to keep their promises to those of promisor dictators asked to keep a promise made by someone else. Clearly, only the former dictators could be driven by MC motivations. Hence, observing a difference in average behavior across the two groups of dictators would provide evidence in favor of MC.[8]

[7] As discussed in the introduction, it is worth noting that Vanberg (2008) provides a different GA test, also used by Kleinknecht (2019). They compare the *Roll* rates of dictators re-matched with recipients involved in a promise to the *Roll* rates of dictators re-matched with recipients who did not receive a promise during the communication phase. Here SOBs are expected to be different (and they are), but the messages are also different. Hence, any potential difference may be explained by the two channels. In any case, we also perform this test and report it in Appendix B.

[8] A fair comparison needs to consider only switched dictators who also made a promise before the switch, as promisors and non-promisors could have different attitude toward *Rolling*. This is the reason why we focus on г-dictators.

Again, we test two hypotheses. We first explore whether subjects are more likely to keep their own word than promises made by others (H4). As in Kleinknecht (2019), we perform Vanberg's test of MC within gender. However, we do it for different levels of SOBs. This leads to four tests to implement Vanberg (2008)'s test combining gender (women/men) and SOBs (high/low) differences. Assuming that H1 holds, the tests are as follows:

**H4 (implication of MC within gender):** For a given switching probability, not-switched Γ-dictators are more likely to *Roll* than switched Γ-dictators who read a promise. That is, $R_i(j, N) > R_i(j, S)$ for $i \in \{M, F\}$ and $j \in \{L, H\}$. Considering all values for $i$ and $j$, we obtain four MC tests.

The next hypothesis is a between-gender test to explore differences in MC sensitivity between genders when SOBs are either high or low. We compare the increase in the average *Roll* rate for men, presumably driven by MC, to the analogous increase for women. In other words, we perform a difference-in-difference comparison in MC sensitivity between men and women.

We run the following tests for each switching probability:

**H5 (different moral sensitivities between genders):** For a given switching probability, male or female non-switched and switched Γ-dictators who read a promise are equally likely to *Roll*. That is $R_F(j, N) - R_F(j, S) = R_M(j, N) - R_M(j, S)$ for $j \in \{L, H\}$.

Note that, H4 is a directional hypothesis, while H5 is not. As in the case of H3, we do not impose any directional hypothesis for H5, as different theories may lead to opposite conclusions.

## 3 Procedures[9]

The experiment was conducted at the CIMEO Experimental Economics Lab of Sapienza University of Rome. The experiment involved 384 undergraduate student subjects (12 sessions of 8 rounds, 32 subjects in each session) recruited using an online system. Upon arrival, subjects were randomly assigned to 32 isolated computer terminals. Three assistants gave instructions (see the Supplementary Material online) and checked that participants followed the procedures correctly. Before starting the experiment, subjects completed a short questionnaire testing their comprehension.

Each session consisted of eight rounds, with a perfect stranger matching. Payoffs, as reported in Fig. 1, were computed in tokens (with one token = 0.50 euro). At the end of each session, one round was randomly chosen for payment. FOBs and SOBs were elicited by asking subjects to guess their counterparts' actions and guesses. Incentives were provided for all rounds except the one chosen for payment, implying

[9] The procedure is the same as in Di Bartolomeo et al. (2019b). We use data from Di Bartolomeo et al. (2019b) augmented with six additional sessions. Hence the gender-based tests have the same statistical power of those in the original study.

that subjects had no incentive to hedge against bad outcomes and, thus, misreport their beliefs.[10] All subjects also received a fixed show-up fee of 2.50 tokens.

In each round, the following five stages were implemented:

1) **Communication.** Subjects were randomly matched to form 16 chatting pairs, with a 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 the communication phase, roles were randomly assigned in each pair, and subjects were informed of that. Then, depending on the treatment, the switching probability was announced as either 25% (low) or 75% (high).
3) **Belief elicitation**. This stage has two parts:

    a. FOBs: each recipient was asked to guess their matched dictator's expected action.
    b. SOBs: dictators were asked to guess the expected action guessed by their matched recipient.

4) **Switching.** Depending on the treatment, 25% or 75% of recipients were switched. Only dictators were informed whether a switch occurred. Dictators with switched recipients were allowed to read the prior conversation of their new recipient.
5) **Dictators' action.** All dictators chose between *Roll* or *Don't Roll*. All subjects were informed of their payoff for the round. Recipients were not informed whether they had been switched, nor could they infer the dictator's choice when their payoffs were zero.[11]

Messages were classified according to Vanberg's protocol. Following Vanberg (2008), we refer to each chat sent by a subject in a round as a "message." We had 3,072 messages (32 subjects in 12 sessions of 8 rounds.) We asked two research assistants to code the messages according to whether they conveyed a promise or statement of intent indicating that the subject would *Roll*. Ex-ante, we decided to randomly use the code of only one of the two assistants. The assistants were unaware of this choice.

We use a within-subject design to compare averages at the session level. Our data are independent at the session level, but not at the individual level.

[10] Our elicitation procedure is described in detail in Appendix.

[11] 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.".

**Table 1** SOBs of switched Γ -dictators

| Treatment | Switching Probability | **Women** (a) | **Men** (b) |
|---|---|---|---|
| (i) | 25% (low) | **0.73** (0.34/266) | **0.76** (0.27/251) |
| (ii) | 75% (high) | **0.66** (0.35/229) | **0.62** (0.35/234) |

**Table 2** *Roll* rates of switched Γ -dictators

| | | **Women** | **Men** |
|---|---|---|---|
| Treatment | Switching Probability | Switch (a) | Switch (b) |
| (i) | 25% (low) [high SOBs] | **0.33** (0.48/45) | **0.50** (0.51/46) |
| (ii) | 75% (high) [low SOBs] | **0.35** (0.48/161) | **0.27** (0.44/161) |

## 4 Results

Our sample contains 2,240 promises out of 3,072 messages. The promise rates of men and women (71% vs. 75%) are not statistically different ($Z=0.47$, $p=0.638$).[12] All statistics are obtained using the Wilcoxon signed-rank test, which compares data at the session level. Our data are independent at the session level, but not at the individual level.[13] We test the hypotheses discussed in Sect. 2.2. We first test for gender differences in SOBs and if our exogenous variation works. We show that there is no difference in SOBs between women and men for any given switching probability (H1a) and that the exogenous variation works for both genders (H1b) in the same manner (H1c).

Table 1 reports the SOBs of Γ-dictators. The switching probabilities are reported by row, and the dictators' gender is listed by column. For instance, the value in the first cell of the table reports that 73% of women-dictators who made a promise believe that the recipient thinks that the dictator will roll after making a promise when the switching probability is low (the average $S_F(L, S) = 0.73$). Standard deviations and number of observations are indicated in parentheses.

We test H1a by comparing the SOBs by row. We do not find any statistically significant difference between men's and women's SOBs, when the chance of being rematched is low, 76% vs. 73% ($Z=1.26$, $p=0.209$), or high, 62% vs. 66% ($Z=-0.16$,

[12] Across all subjects, the frequency of promises does not statistically differ across high/low switching probability treatments (i.e., 76% vs. 73%: $Z=1.16$, $p=0.247$).

[13] We use one-tailed (two-tailed) tests when we do (do not) have preconceived directional hypotheses. Note that data reported in the tables are all sample averages.

**Table 3** *Roll* rates of all Γ -dictators

| Treatment | Switching Probability | Women | | Men | |
|---|---|---|---|---|---|
| | | Switch (a) | No-switch (b) | Switch (c) | No-switch (d) |
| (i) | 25% (low) [high SOBs] | **0.33** (0.48/45) | **0.52** (0.50/221) | **0.50** (0.51/46) | **0.53** (0.50/205) |
| (ii) | 75% (high) [low SOBs] | **0.35** (0.48/161) | **0.56** (0.50/68) | **0.27** (0.44/161) | **0.60** (0.49/73) |

$p=0.875$). Our results support the non-existence of a gender difference in SOBs (H1a).

The data in Table 1 are also consistent with exogenous variations in expectations (H1b). Promisors' SOBs reported in column (a) are high (low) when the chance of being re-matched is low (high). The result holds for each gender: for women (column (a)): 0.73 vs. 0.66 ($Z=2.04$, $p=0.020$); and for men (column (b)): 0.76 vs. 0.62: ($Z=3.06$, $p=0.001$).[14]

Finally, the data supports H1c. The observed increase in SOBs for men is 0.14, while for women is 0.07 and 0.14 vs. 0.07 are not statistically different ($Z=1.17$, $p=0.239$). We do not find any significant differential movement in second-order beliefs, but the directional effect indicates that men change their beliefs more strongly than women. This is a potential concern for the identification of GA, as we will find that men are more sensitive to GA than women are.[15] Alternative approaches can fix second-order beliefs by design and thus do not require the assumption that these beliefs are differentially affected by the treatments (Bellemare et al., 2018, 2019). Here, we follow Vanberg's approach (2008) for comparison with related literature.[16]

We focus next on GA (H2-H3). Table 2 reports the *Roll* rates of switched Γ-dictators when the probability of switching is low or high. As we consider only switched pairs, dictators' motivations cannot be related to MC. Moreover, switched dictators' SOBs are likely to be high when the switching probability is low and vice versa.

Remember that Γ-dictators are matched with recipients who received a promise. Again, switching probabilities are reported by row, and the dictators' gender is listed by column.

For instance, the value in the first cell of Table 2, 0.33, is the average *Roll* rate of female switched Γ-dictators playing the game with a low switching probability (and high SOBs), i.e., $R_F(L, S)$. Again, standard deviations and number of observations are indicated in parentheses.

[14] Here, we have a preconceived directional hypothesis and use a one-tailed test. However, exogenous variations are still significant at the 5% significance level when using the most conservative two-tailed test.

[15] It is interesting to note that Kleinknecht (2019) finds that promises raise the second-order beliefs of women more so than for men (hence, the opposite pattern).

[16] A detailed discussion of pros and cons of the two approaches is beyond the scope of the present paper.

Next, we look at H2 (GA) for men and women. The average *Roll* rate of men is significantly higher when the probability of being re-matched is low (high SOBs) rather than high (low SOBs): 0.50 vs. 0.27 ($Z=2.24$, $p=0.013$). Conversely, the average *Roll* rate of switched women is not significantly different in the two matching probability treatments: 0.33 vs. 0.35 ($Z=-0.20$, $p=0.422$). That is, H2 is confirmed for men but not for women. This supports the idea that GA drives men's motivations to a greater extent than women's.[17] We further explore this point in H3.

To test H3 (guilt aversion sensitivities), we compare the change in *Roll* rates of men for the two different matching probabilities (0.23 = 0.50–0.27) to the corresponding change in *Roll* rates of women (-0.02 = 0.33–0.35). H3 is supported: 0.23 vs. -0.02 ($Z=1.73$, $p=0.084$).[18]

We finally test for MC (H4-H5). We augment Table 2 with additional information about the *Roll* rate of not-switched Γ-dictators.

Table 3 reports the *Roll* rates of switched and non-switched Γ-dictators by gender (columns) for the different switching probabilities (rows). Remember that SOBs are likely to be high when the switching probability is low and vice versa. Again, standard deviations and number of observations are indicated in parentheses.

Following Vanberg (2008), we test for MC by comparing the average *Roll* rate of non-switched dictators (who made a promise) to that of switched dictators (who made a promise and are re-matched with a recipient who received a promise from someone else). We test MC motivation within each gender (H4).

We run four tests by combining the two genders and the two switching probabilities (i.e., two levels of SOBs.) H4a (H4b) focuses on women's MC-driven motivation when the switching probability is low (high), while H4c (H4d) focuses on men's MC-driven motivation when the switching probability is low (high).

The outcomes of the tests are listed below (all tests are two-tailed.)

1. H4a: $R_F(L,N) > R_F(L,S)$ 0.52 vs. 0.33 ($Z=2.39$, $p=0.016$)
2. H4b: $R_F(H,N) > R_F(H,S)$ 0.56 vs. 0.35 ($Z=2.31$, $p=0.020$)
3. H4c: $R_M(L,N) > R_M(L,S)$ 0.53 vs. 0.50 ($Z=0.08$, $p=0.937$)
4. H4d: $R_M(H,N) > R_M(H,S)$ 0.60 vs. 0.27 ($Z=2.12$, $p=0.034$)

Our results support MC-driven motivations for women (H4a and H4b), while MC-driven motivations for men are only supported when SOBs are low (H4d).

Finally, we test for MC sensitivity between genders (H5). We have two cases, one for each SOB level (H5a and H5b for low and high SOBs, respectively.) We look at gender differences by rows.

Let us start with the case of low SOBs; if the switching probability is high (row (ii) in Table 3), the data do not provide support for any difference in the MC impact within genders, i.e., (0.56–0.35 = 0.21) vs. (0.60–0.27 = 0.33), $Z=-0.18$, $p=0.859$. Hence, H5a holds.

[17] In this case, we have a preconceived directional hypothesis and use a one-tailed test. However, our result (GA for men) is still significant at a 5% significance level even if we use the most conservative two-tailed test.

[18] We use a two-tailed test, the result is weakly significant at a 10% level.

Next, we focus on the case of high SOBs (row (i) in Table 3). We compare the change in *Roll* rates of women (0.52–0.33 = 0.19) to the corresponding change in *Roll* rates of men (0.53–0.50 = 0.03). If SOBs are high, H5b is not supported at the 5%-level although it is close: compare 0.19 vs. 0.03 ($Z = 1.73$, $p = 0.084$).[19]

Our results tend to support the conclusion that women are sensitive to MC independently of their SOBs, while men are only sensitive to MC when their SOBs are low.

## 5 Conclusion

We explored guilt aversion (GA) and moral commitment (MC) motivations in a random dictator game with pre-play communication. With the main exception of Kleinknecht's (2019) contribution discussed above, few studies focused on GA- or MC-related gender effects and the results are mixed. Our experimental results support the idea that moral commitment drives women's behavior more than men's and the hypothesis that men are more guilt-averse than women.

We discussed our GA-finding with the psychologist Leda Cosmides, who, along with her husband, anthropologist John Tooby, made seminal contributions to the development of evolutionary psychology. Tooby and Cosmides (2008, p. 177) argue that "each emotion evolved to deal with a particular, evolutionarily recurrent situation." This statement may suggest that emotions, such as guilt, may affect men and women differently if the games they play cast the sexes in asymmetric roles.

Our finding is then consistent with the following perspective: men may have a more substantial evolutionary motivation to be known to be prone to guilt for child-bearing reasons. A pregnant mother spends nine-month gestation when the father could conceivably take off and produce offspring with other women. If women anticipated such opportunistic behavior, they might not agree to conceive in the first place. If a prospective father is sensitive to feelings of guilt, that may prevent him from leaving the mother and, therefore, secure her trust.[20]

Granted, other situations may favor women's sensitivity to guilt more than that of men. When we presented the story of the previous paragraph to Leda Cosmides, she suggested one: A man who is investing in a woman's offspring (post-birth) must trust that she has been sexually faithful to him. In this situation, she may benefit more than he does from being guilt prone. All in all, this may invite the reflection that gender differences could be domain specific.

[19] A conservative two-tailed test is used.

[20] For another story favoring male guilt, unrelated to pregnancy, see Dufwenberg's (2002) analysis of a "marital investment game" featuring a trusting wife with a guilt-averse husband: The "asymmetric treatment of the sexes is consistent with Weitzman's (1986, p. 67) observation that: '[h]usbands and wives typically invest in careers—most particularly in the husband's education and career—and the products of such investments are often a family's major assets,' with Borenstein and Courant's (1989; Footnote 3) observation that a medical student with a supporting spouse typically is a husband with a wife, with evidence concerning divorce cases decided in U.S. courts (Polsby and Zelder 1994, Footnote 4), and with Cohen's (1987) general finding that nuptial gains tend to accrue to men early on in a marriage and to women toward the end.".

**Table 4** 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 |

**Table 5** *Roll* rates of Γ -dictators.

| Treatment | Switching probability | Message | Gender | |
|---|---|---|---|---|
| | | | Women (a) | Men (c) |
| (i) | 25% | Read (high SOBs) | **0.33** | **0.50** |
| | | | (0.48/45) | (0.51/46) |
| | | Do not Read (low SOBs) | **0.18** | **0.22** |
| | | | (0.39/17) | (0.42/23) |
| (ii) | 75% | Read (high SOBs) | **0.35** | **0.27** |
| | | | (0.48/161) | (0.44/161) |
| | | Do not Read (low SOBs) | **0.30** | **0.29** |
| | | | (0.46/50) | (0.46/56) |

However, our lab experiment resembles neither of the two evolutionarily potentially relevant settings we just described. The idea should be properly tested in an appropriate experimental design that captures the evolutionary elements. This is beyond the scope of the current paper and, therefore, left to future developments, which we believe are promising.

## Appendix A

This appendix describes the FOBs and SOBs elicitation procedure. After the communication phase, recipients were asked to guess what their (unknown) dictator would choose to do. They had been told the switching probability (either 75% or 25%). Recipients expressed their guesses by ticking one of the five-point scale in Table 4. This scale is the same as in Vanberg. Beliefs are then re-scaled to 1.00, 0.75, 0.50, 0.25, and 0.00.

After dictators were told whether their paired recipient had been switched and read their previous communication, they were asked to guess their guess. Specifically, they had to guess which of the five points of Table 4 had been ticked by their counterpart. Correct guesses earned 0.50 tokens.

## Appendix B

The paper focuses on between and within gender clean tests of GA and MC for Γ -dictators. Here, we perform four additional tests akin to Vanberg's (2008) and Kleinknecht's (2019) GA test. Recipients who received a promise hold higher FOBs than recipients who did not receive a promise. GA test can then be built by comparing the behavior of dictators whose switched recipients received a promise from someone else with that of dictators whose switched recipients did not receive a promise.

Our data are reported in Table 5, where we also consider the *Roll* rates of switched Γ-dictators matched with a recipient who did not receive a promise.

Table 5 allows us to compare the behavior of switched Γ-dictators facing a recipient who, everything else equal, received a promise (high SOB) to one who did not (low SOB). However, as discussed in the main text, the two kinds of compared dictators differ along two dimensions: 1) SOBs and 2) message read.) Hence, the outcomes of these tests in supporting or not GA can be questionable.

Our findings are as follows:

1. In treatment (i), the data support GA or a different attitude toward the message read between men and women. For the latter, we compare 0.33 (high SOBs) vs. 0.18 (Low SOBs): $Z = 2.34$, $p = 0.019$, while for men, we compare 0.50 (high SOBs) vs. 0.22 (low SOBs): $Z = 2.30$, $p = 0.022$. Differently from Kleinknecht (2019), we find support for men GA.
2. In treatment (ii), no GA support is found. In the case of the women, we compare 0.35 (high SOBs) vs. 0.30 (low SOBs): $Z = 1.06$, $p = 0.289$. In the case of men, we compare 0.27 (high SOBs) vs. 0.29 (low SOBs): $Z = 0.94$, $p = 0.347$.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s40881-024-00180-8.

**Acknowledgements** We are grateful to Giuseppe Attanasi and Leda Cosmides for helpful discussions. Project funded by the European Union – Next Generation EU (UPB SpagnoloG22 Prin CUP: E53D23006140006).

**Funding** Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement.

**Data availability** Further information and data can be requested directly from the authors.

## References

Attanasi, G., Rimbaud, C., & Villeval, M. C. (2019). Embezzlement and guilt aversion. *Journal of Economic Behavior & Organization, 167*, 409–429.

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

Battigalli, P., & Dufwenberg, M. (2022). Belief-dependent motivations and psychological game theory. *Journal of Economic Literature, 60*, 833–882.

Bellemare, C., & Sebald, A. (2023). Measuring Belief-dependent Preferences without Data on Beliefs. *The Review of Economic Studies, 90*, 40–64.

Bellemare, C., Sebald, A., & Suetens, S. (2017). A note on testing guilt aversion. *Games and Economic Behavior, 102*, 233–239.

Bellemare, C., Sebald, A., & Suetens, S. (2018). Heterogeneous Guilt Sensitivities and Incentive Effects. *Experimental Economics, 21*, 316–336.

Bellemare, C., Sebald, A., & Suetens, S. (2019). Guilt Aversion in Economics and Psychology. *Journal of Economic Psychology, 73*, 52–59.

Benetti-McQuoid, J., & Bursik, K. (2005). Individual differences in experiences of and responses to guilt and shame: Examining the lenses of gender and gender role. *Sex Roles: A Journal of Research, 53*(1–2), 133–142.

Bolton, G. E., & Katok, E. (1995). An experimental test for gender differences in beneficent behavior. *Economics Letters, 48*, 287–292.

Borenstein, S., & Courant, P. N. (1989). How to carve a medical degree: Human capital assets in divorce settlements. *American Economic Review, 79*, 992–1009.

Cartwright, E. (2019). A survey of belief-based guilt aversion in trust and dictator games. *Journal of Economic Behavior & Organization, 167*, 430–444.

Charness, G., Di Bartolomeo, G., and S. Papa (2023), “A stranger in a strange land: Promises and identity,” *Games and Economic Behavior,* forthcoming.

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

Childs, J. (2012). (2012), “Gender differences in lying.” *Economics Letters, 114*, 147–149.

Ciccarone, G., Di Bartolomeo, G., & Papa, S. (2020). The rationale of in-group favoritism: An experimental test of three explanations. *Games and Economic Behavior, 124*, 554–568.

Cohen, L. (1987). Marriage, divorce, and quasi rents; or, I gave him the best years of my life. *Journal of Legal Studies, 16*, 267–303.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature, 47*, 448–474.

Dhami, S., Wei, M., & A. al-Nowaihi,. (2019). Public goods games and psychological utility: Theory and evidence. *Journal of Economic Behavior and Organization, 167*, 361–390.

Di Bartolomeo, G., Dufwenberg, M., & Papa, S. (2019a). The sound of silence: A license to be selfish. *Economics Letters, 182*, 68–70.

Di Bartolomeo, G., Dufwenberg, M., & Papa, S. (2023a). Promises and partner-switch. *Journal of the Economic Science Association, 9*, 77–89.

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

Di Bartolomeo, G., Dufwenberg, M., Papa, S., & Passarelli, F. (2023b). Promises or agreements? *Moral Commitments in Bilateral Communication, Economics Letters, 222*(1), 4.

Dreber, A., & Johannesson, M. (2008). Gender differences in deception. *Economics Letters, 99*, 197–199.

Dufwenberg, M. (2002). Marital investments, time consistency and emotions. *Journal of Economic Behavior & Organization, 48*, 57–49.

Dufwenberg, M., & Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games & Economic Behavior, 30*, 163–182.

Eckel, C. C., & Grossman, P. J. (1998). Are women less selfish than men? Evidence from dictator experiments. *Economic Journal, 108*, 726–735.

Ellingsen, T., & M., Johannesson, S., Tjøtta, and G. Torsvik,. (2010). Testing guilt aversion. *Games and Economic Behavior, 68*, 95–107.

Else-Quest, N. M., Higgins, A., Allison, C., & Morton, L. C. (2012). Gender differences in self-conscious emotional experience: A meta-analysis. *Psychological Bulletin, 138*(5), 947.

Ferguson, T. J., & Crowley, S. L. (1997). Gender differences in the organization of guilt and shame. *Sex Roles, 37*, 19–44.

Kleinknecht, J. (2019). A man of his word? An experiment on gender differences in promise keeping. *Journal of Economic Behavior & Organization, 168*, 251–268.
Nihonsugi, T., Tanaka, T., & Haruno, M. (2022). Gender differences in guilt Aversion in Korea and the United Kingdom. *Scientific Reports, 12*, 8187.
Polsby, D. D., & Zelder, M. (1994). Risk-adjusted valuation of professional degrees in divorce. *Journal of Legal Studies, 33*, 273–285.
Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 114–137). The Guilford Press.
Vanberg, C. (2008). Why do people keep their promises? An experimental test of two explanations. *Econometrica, 76*, 1467–1480.
Weitzman, L. J. (1986). A new look at career assets. *Ms. Magazine., 24*, 67–68.