

# Guilt Aversion: Eve versus Adam\*

Giovanni Di Bartolomeo,<sup>a</sup> Martin Dufwenberg,<sup>b</sup> Stefano Papa,<sup>c</sup> and Laura Razzolini<sup>d</sup>

May 19, 2022

**Abstract.** Our study contributes to a large literature in experimental economics that explores gender differences in how people are motivated. We focus on guilt aversion (GA), a surprisingly rather unexplored issue. Our experiment supports the idea that men are more GA than women. Our results also support different rationales to explain observed similar behaviors, like promise keeping. We provide a potential intuition for our findings, which is based on the pregnancy-related biological asymmetry between genders.

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

**Keywords:** gender, guilt aversion, promises, evolutionary psychology

## 1. Introduction

Evolutionary psychologists Tooby and Cosmides (2008, p. 177) argue that “each emotion evolved to deal with a particular, evolutionarily recurrent situation.” This may suggest that emotions affect men and women differently, if the games they play cast the sexes in asymmetric roles. A case in hand may concern *guilt*. Men may have stronger evolutionary reasons to be known to be prone to guilt for child-bearing reasons. A pregnant mother spends nine months in gestation, during which time 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 known to be prone to feeling guilty that may prevent him from leaving the mother, and therefore help secure her trust.

Contributions in behavioral and experimental economics have developed a theory of guilt aversion (GA) in games and show that it can help explain trustworthy behavior in the lab.<sup>1</sup> The focus of this

---

\* We are grateful to Leda Cosmides for a helpful discussion.

<sup>a</sup> Department of Economics and Law, Sapienza University of Rome; CIMEO; Department of Economics, University of Antwerp.

<sup>b</sup> Department of Economics, University of Arizona; University of Gothenburg; CESifo; [martind@eller.arizona.edu](mailto:martind@eller.arizona.edu)

<sup>c</sup> Department of Economics and Finance, University of Rome Tor Vergata; CIMEO.

<sup>d</sup> Department of Economics, Finance and Legal Studies, University of Alabama.

literature has not concerned with gender differences. Motivated by the trust-during-pregnancy example given above, in this paper, we experimentally explore whether GA affects men more than women.

The directional nature of such an effect is not self-evident. While the nine-months gestation story, given above, seems important, other stories may be provided in support of female guilt-proneness. We return to and discuss this topic further in the final section, after reporting our data.

The general formulation of GA theory, due to Battigalli and Dufwenberg (2007), involves a form of belief-dependent motivation. This poses experimental design challenges, concerning how to induce exogenous variations of key beliefs. Di Bartolomeo *et al.* (2019), expanding on ideas pioneered by Vanberg (2008), devised techniques designed to overcome the issue. We rely on these methods as well. Section 2 explains in more detail our design. We also draw connections to a recent study by Kleinknecht (2019) which used Vanberg's design to study gender differences, but did not adopt Di Bartolomeo *et al.*'s techniques that allow for a more nuanced focus on GA.

Sections 3 and 4 present our results and Section 5 offers a concluding discussion.

## 2. The experiment

### The game

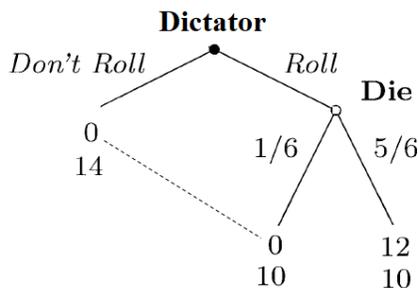
Following Vanberg, we use a two-player mini-dictator game augmented with a) bilateral pre-play communication; b) roles' randomness and asymmetric information; and c) a partner-switching mechanism. It works in five steps:

- 1) Participants are matched in pairs. Each pair can communicate by sending messages. The subjects know that they will play a mini-dictator game, but not who will act in which role. Subjects can make promises about their behavior in case they are chosen to be dictators.
- 2) After the communication phase, each player is assigned to a role, either dictator or recipient.
- 3) After the communication phase and before making the allocation decision, a given proportion of dictators have their partners switched. The proportion is known by both players from the beginning, 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 received by their new partners.

---

<sup>1</sup> Battigalli and Dufwenberg (2007) develop a general theory. Dufwenberg and Gneezy (2000) and Charness and Dufwenberg (2006) ran early experiments. Cartwright (2019) surveys the subsequent experimental literature. See Battigalli and Dufwenberg (2021, especially Sections 3.1 and 7) for a broad related discussion.

- 5) Each pair plays the game (form) shown in Figure 1. Dictators choose between *Roll* and *Don't Roll*. The payoffs are in Euro and the recipient's payoff is listed on top; the information set "0...0" indicates that the recipient is not informed how the payoff of "0" arose).



**Figure 1 – Game payoffs**

### Treatments

Following Di Bartolomeo *et al.* (2019a), we run two treatments, by varying the switching probability, with a low probability of 1/4 or a high switching probability of 3/4. In the original design, Vanberg (2008) used a switching probability of 1/2. He hypothesized that dictators who promise to *Roll* have a preference to keep their own word. Using the 1/2 switching probability, Vanberg confirmed the hypothesis that subjects who promise to *Roll* are more likely to *Roll* if their recipient is not switched. A guilt averse dictator who chooses *Don't Roll* suffers guilt proportional to his “second-order belief” (SOB) of how strongly he believes the recipient “first-order” believes (FOB) that he will choose to *Roll*.<sup>2</sup> Dufwenberg and Gneezy (2000) and Charness and Dufwenberg (2006) (in different, but comparable games) report support for GA via the hypothesis that the higher is a player’s SOB, the more likely he is to make a pro-social choice, like *Roll*. Conclusions based on subjects’ beliefs, however, are questionable, because the observed correlations do not necessary imply causation. Several studies have proposed new methods to deal with this issue.<sup>3</sup> Di Bartolomeo *et al.* (2019a) realized that dictators who issued a promise to *Roll*, should *Roll* if the switching probability is low (since fewer subjects would be switched). Recipients would presumably anticipate that, and so hold corresponding beliefs, independently of whether or not they were actually switched (since recipients are not told whether there was indeed a switch). Dictators would

<sup>2</sup> More precisely, if  $p$  is the probability with which the recipient believes the dictator will choose *Roll*, then the dictator’s SOB is the mean of his probability measure over values of  $p$ .

<sup>3</sup> See, e.g., Ellingsen *et al.* (2010), Khalmetski *et al.* (2015), Bellemare *et al.* (2017), Ederer and Stremitzer (2017), Attanasi *et al.* (2019), Dhimi *et al.* (2019), and Di Bartolomeo *et al.* (2019, 2020). Charness and Dufwenberg pointed to the problem but didn’t address it in their design. For a broader discussion – contrasting some of these approaches, highlighting some controversies, and also drawing connections to parallel issues surrounding Vanberg’s (2008) work on promise-keeping – see Battigalli and Dufwenberg (2021, Section 7).

presumably anticipate that, and hold the corresponding SOBs. This leads to the following critical test for GA: *If the switching probability is low, since SOBs are high, dictators should be more likely to Roll.* We indeed test for GA using two treatments which differ by the value of the switching probability (1/4 or 3/4). We also account for the gender of dictators and recipients, thus testing for gender differences in GA.

### **Justification relative to previous work**

Charness and Dufwenberg's (2006) main focus concerned articulating and testing a GA-based theory for how promises may foster trust and cooperation. The idea is that promises raise co-player's FOB and a promisor's SOB, that the promise will be fulfilled, and GA makes this credible and self-fulfilling. Vanberg (2008) presented his story as an alternative to Charness and Dufwenberg's (2006) (see Di Bartolomeo *et al.* (2019, 2021) for more details and in-depth discussion).

While Vanberg's (2008) primary focus was his own theory of promise-keeping, his approach nevertheless permits the following test of GA: Focusing on switched dictators who sent a promise, he explored how their behavior changed depending on whether or not their new recipient had previously received a promise from someone else. He tested in this way for GA, because recipients who received a promise hold higher FOBs than recipients who did not receive a promise. Vanberg did not find support for GA. Kleinknecht (2019), however, in a study focused on gender differences, replicated Vanberg's design and did find support for women's GA, i.e., women are likely to keep their promises to not let the other's beliefs down.<sup>4</sup>

Vanberg's and Kleinknecht's tests have the following questionable feature: In their studies, they compared the behavior of dictators, whose switched recipients received a promise from someone else, with the behavior of dictators whose switched recipients did not receive a promise. Such a comparison involves subjects who differ in two ways. First, recipients who received a promise have higher FOBs than those who did not. Second, the comparison involves subjects who received a promise and recipients who did not. A comparison along two dimensions (rather than one) is undesirable from an experimental methodology point of view. The design by Di Bartolomeo *et al.* (2019) avoids this problem, since in their treatment, the comparison involves subjects who differ only in their expectation and not in whether or not they received a promise. Di Bartolomeo *et al.*, like Vanberg, did not find support for GA. However, unlike Kleinknecht, Di Bartolomeo *et al.* did not record subjects' gender. Our current design fills the gap, using two treatments with different switching probabilities and accounting for gender effects.

---

<sup>4</sup> It is worth noting that Kleinknecht (2019) results are based on first round observations only. By contrast, Vanberg (2008) and Di Bartolomeo *et al.* (2019) use session level comparisons on all the sample.

## Hypotheses

We focus on dictators who both made and “read” promises, and refer to this group as the  $\Gamma$ -dictators. Non-switched  $\Gamma$ -dictators are those who must decide whether to keep their own promises, while switched  $\Gamma$ -dictators are promisors who are rematched with a co-player who received a promise from another dictator. As recipients are not informed whether a switch occurred, it follows that FOBs and SOBs are independent of that. 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.

We test two main hypotheses. The first one concerns the following implication of GA:

**H1 (implication of GA):** Switched  $\Gamma$ -dictators (those who issued promises and “read” promises) are more likely to *Roll* when the switching probability is low.<sup>5</sup>

The logic of H1 is that since a  $\Gamma$ -dictator’s SOB is higher when the switching probability is low, then if said  $\Gamma$ -dictator is guilt averse, he is more likely to *Roll*. We test H1 separately for men and women.

Our second main hypothesis explores whether men and women have different guilt sensitivities. If this were the case then, for a given switching probability (and hence SOB’s of a  $\Gamma$ -dictator), the gender with the highest guilt-sensitivity would be more likely to *Roll*:

**H2 (different guilt sensitivities for men and women):** Male and female switched  $\Gamma$ -dictators are not equally likely to *Roll*, for a given switching probability.

Consistently with the findings of Di Bartolomeo *et al.* (2019a), H1 and H2 focus on switched dictators to test for gender differences in GA. In addition, we will also test the robustness of the results of Di Bartolomeo *et al.* (2019a) on promise keeping and the possibility of gender differences in this domain. That is, we will test if men and women have different attitudes in promise keeping. We shall also run replications of Vanberg’s test for GA, as described above (including the “questionable feature”).

---

<sup>5</sup> A fair comparison indeed needs to consider only switched dictators who also made a promise before the switch, as promisor and non-promisors could have different attitude towards *Rolling* and could be in different proportions in the high and low switching groups that are being compared.

### 3. Main results

#### Preliminaries

Our sample consists of 3,072 messages, 1,520 written by women and 1,552 written by men. Women made 1,140 promises to *Roll*, while men issued 1,100 promises.<sup>6</sup> The promise rates made by men and women (71% vs. 75%) are not statistically different ( $Z=0.47$ ,  $p=0.638$ ). All 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.<sup>7</sup>

Before moving to H1 and H2, we check if our design leads to exogenous variations in expectations as intended. Our sample consists of 1,536 dictators' decisions. SOBs are reported in Table 1. We report SOBs for the subset of the dictators who made a promise during the communication stage and who, after the eventual switch, played with a recipient who also received a promise from someone else. Standard deviations and number of observations are indicated in parentheses.

The data in Table 1 are consistent with exogenous variation in expectations: Promisors' SOBs reported in column (a) are high (low) when the chance of being re-matched is low (high): 0.75 is significantly higher than 0.64 ( $Z=2.75$ ,  $p=0.006$ ). The same result holds for the two genders: for women (column (b)): 0.73 vs. 0.66 ( $Z=2.04$ ,  $p=0.04$ ); and for men (column (c)) 0.76 vs. 0.62: ( $Z=3.06$ ,  $p=0.002$ ).

Moreover, comparing the SOBs in each row, we do not find any statistically significant difference between men and women when the chance of being re-matched is low, 76% vs. 73% ( $Z=1.26$ ,  $p=0.209$ ) or high, 62% vs. 66% ( $Z=-0.16$ ,  $p=0.875$ ). This supports non-existence of a gender difference in SOBs.

**Table 1 – SOBs of  $\Gamma$ -dictators**

TREATMENT	SWITCHING PROBABILITY	All Dictators (a)	Women (b)	Men (c)
(i)	25% (low)	<b>0.75</b> (0.31/517)	<b>0.73</b> (0.34/266)	<b>0.76</b> (0.27/251)
(ii)	75% (high)	<b>0.64</b> (0.35/463)	<b>0.66</b> (0.35/229)	<b>0.62</b> (0.35/234)

Table 2 reports the average *Roll* rates of the dictators who are asked to keep promises made by others (switched) and of those who are asked to keep their own promises (non-switched). Both are distinguished by gender and switching probabilities.

<sup>6</sup> Messages were classified by a research assistant and coded according to whether they conveyed a promise to *Roll*.

<sup>7</sup> In comparing men versus women, we use two-tailed tests as we do not have preconceived directional hypotheses. Instead, at the end of the session, when we test the rationale of promise keeping, we use one-tailed tests since we have directional hypotheses.

**Table 2 – Roll rates of  $\Gamma$ -dictators**

TREATMENT	SWITCHING PROBABILITY	Women		Men	
		SWITCH (a)	NO-SWITCH (b)	SWITCH (c)	NO-SWITCH (d)
(i)	25% (low) [high SOBs]	<b>0.33</b> (0.48/45)	<b>0.52</b> (0.50/221)	<b>0.50</b> (0.51/46)	<b>0.53</b> (0.50/205)
(ii)	75% (high) [low SOBs]	<b>0.35</b> (0.48/161)	<b>0.56</b> (0.50/68)	<b>0.27</b> (0.44/161)	<b>0.60</b> (0.49/73)

## H1 and H2

Testing for GA via H1 requires looking at the behavior of switched dictators (columns (a) and (c)) and comparing cells by rows. The average *Roll* rate of men is significantly higher when the chance of being re-matched is low (high SOBs) rather than high (low SOBs): 0.50 vs. 0.27:  $Z=2.24, p=0.025$ . Conversely, the average *Roll* rate of switched women is not significantly different when the chance of being re-matched is different: 0.35 vs. 0.33:  $Z=0.26, p=0.398$ . This result supports the idea that men (women) are more (less) likely to keep the promises made by others when the switching probability is low; that is, we find evidence in support of men being guilt averse, but do not find support that women are guilt averse.

Looking at the behavior of switched dictators (columns (a) and (c)), we test H2 by comparing cells across columns. The average *Roll* rate of men is significantly higher than that of women when the chance of being re-matched is low (high SOBs): 0.50 vs. 0.33:  $Z=1.96, p=0.049$ . In contrast, the average *Roll* rate of men is not significantly higher than that of women when the chance of being re-matched is high (low SOBs): 0.27 vs. 0.35:  $Z=-0.75, p=0.456$ . This result supports the idea that, when SOBs are high (low), men’ guilt sensitivity is (is not) higher than women’ guilt sensitivity.

## 4. Additional findings

Although the focus of the paper is GA, we also test the robustness of the related results from Di Bartolomeo *et al.* (2019a) on the rationale of promise keeping and eventual gender differences.<sup>8</sup> In our setup,<sup>9</sup> we test whether there are gender differences in keeping own promises as compared to keeping

<sup>8</sup> Not surprising, our experiment also provides evidence for promise keeping results, i.e., non-switched dictators are more likely to *Roll* when they made promises as compared to the case when they did not. Results are available upon request.

<sup>9</sup> Here communication is repeated and multi directional and the game structure would eventually support a sort of explicit or implicit conditional agreement: “I promise to *Roll* in the case I will be the dictator if you promise to *Roll* in the same situation.” Indeed, it is worth noting that promise keeping motivations can be sensible to the kind of game and the kind of communication considered (see, e.g., Di Bartolomeo *et al.*, 2019a; 2022).

promises made by others. For each switching probability, we can test if dictators are more likely to keep their own words compared to keeping promises made by others. That is, we can test if people have a preference to keep their own word as suggested for example by Vanberg (2006), controlling for different levels of beliefs.

Focusing on women, our finding strongly supports the idea that the value a person attaches to keeping own promises is independent of own beliefs. That is, women have a preference to keep their own word: In Table 2, compare in the high beliefs case: 0.52 vs. 0.33:  $Z=2.394, p=0.016$ ; and in the case of low beliefs: 0.56 vs. 0.35:  $Z= 2.312, p=0.020$ .) By contrast, men’s GA offers a different picture. In the case of high switching probability and low SOBs, men are more likely to keep their own promises, as are women (0.60 vs. 0.27:  $Z= 2.118, p=0.034$ ), but when SOBs are high, men are equally likely to keep their own promises as they are to keep the promises made by someone else (0.50 vs. 0.53:  $Z=- 0.078, p=0.937$ ).

To obtain clean variation in expectations, we used different switching probabilities. Following Vanberg (2008), GA could be also tested by focusing on the behavior of switched dictators for a given switching probability and how such behavior changes depending on whether or not the new recipient had previously received a promise from someone else (Vanberg’s GA-test). Formally, given a switching probability, Vanberg’s GA-test requires comparing the behavior of switched  $\Gamma$ -dictators playing with a recipient who received a promise from another dictator to the behavior of switched  $\Gamma$ -dictators playing with a recipient who did not receive a promise from another dictator. This comparison is not exactly fair (or genuine), as dictators differ along two dimensions: expectations and the message read. Still, we report what we find.

The test can be performed augmenting Table 2 data with information about the *Roll* rates of switched  $\Gamma$ -dictators matched with a recipient who did not receive a promise. This is reported in Table 3.

**Table 3 – *Roll* rates of  $\Gamma$ -dictators**

TREATMENT	SWITCHING PROBABILITY	MESSAGE	Gender	
			WOMEN (a)	MEN (c)
(i)	25%	Read (high SOBs)	<b>0.33</b> (0.48/45)	<b>0.50</b> (0.51/46)
		Do not Read (low SOBs)	<b>0.18</b> (0.39/17)	<b>0.22</b> (0.42/23)
(ii)	75%	Read (high SOBs)	<b>0.35</b> (0.48/161)	<b>0.27</b> (0.44/161)
		Do not Read (low SOBs)	<b>0.30</b> (0.46/50)	<b>0.29</b> (0.46/56)

Table 3 allow performing four tests (akin to Vanberg’s GA-test), all comparing the behavior of switched  $\Gamma$ -dictators facing a recipient who, everything else equal, received a promise (high SOB) to one who did not (low SOB). In treatment (i), where SOBs are relatively high, the data support GA (or a different attitude towards the message read) by men and women: 0.33 (High SOB) vs. 0.18 (Low SOB sob):  $Z=2.34, p=0.019$  and 0.50 (High SOB) vs. 0.22 (Low SOB):  $Z=2.30, p=0.022$ . Instead, in treatment (ii), where SOBs are relatively low, no GA support is found for both women and men: 0.35 (High SOB) vs. 0.30 (Low SOB):  $Z=1.06, p=0.289$  and 0.27 (High SOB) vs. 0.29 (Low SOB):  $Z=0.94, p=0.347$ .

Our as well as Vanberg’s GA-tests can be also performed by focusing on the set of dictators who did not made promises. Their *Roll* rates are reported in Table 4. It is worth noting that these tests are less robust than those previously done because of the relatively lower number of observations.

**Table 4 – *Roll* rates of dictators who did not promises**

TREATMENT	SWITCHING PROBABILITY	MESSAGE	Gender	
			WOMEN (a)	MEN (c)
(i)	25%	Read (high SOBs)	<b>0.43</b> (0.51/23)	<b>0.24</b> (0.44/17)
		Do not Read (low SOBs)	<b>0.50</b> (0.55/6)	<b>0.47</b> (0.52/15)
(ii)	75%	Read (high SOBs)	<b>0.24</b> (0.43/49)	<b>0.26</b> (0.57/44)
		Do not Read (low SOBs)	<b>0.20</b> (0.41/20)	<b>0.27</b> (0.46/22)

Performing this same test using the data of Table 2 (**our tests**) deliver opposite results. Women are GA, while men are not. In Table 4, we consider male and female dictators who did not make a promise, while the dictators considered in Table 2 all made a promise. We compare now the behavior (*Roll* rates) of switched dictators in the low-switch-probability treatment who did not make a promise and played with a recipient who received a promise from another dictator with the behavior of switched dictators in the high-switch-probability treatment who did not make a promise and played with a recipient who received a promise from another dictator. For men, data do not support GA: 0.24 vs. 0.26:  $Z=0.14, p=0.888$ ; while for women they do: 0.43 vs. 0.24:  $Z=2.19, p=0.028$ .

As the action of making or not a promise is endogenous, the above results combined with those from Table 2 suggest a sort of self-selection. Assuming that some individuals are guilt averse and others are not guilt averse, then guilt averse men (women) are more (less) likely to make promises. In other words, male promisors are more prone to GA, while men who did not promise are not. Conversely, female promisors seem less prone to GA, while women who did not promise appear not to be so.

We can finally replicate Vanberg’s GA-test also for those dictators who did not make a promise. In such a case, no support for GA is found (or for a different attitude regarding the messages read).<sup>10</sup> Formally, in Table 4, treatment (i), where SOBs are relatively high, the data do not support GA for both men and women: 0.24 (High SOB) vs. 0.47 (Low SOB):  $Z=-0.65$ ,  $p=0.517$  and 0.43 (High SOB) vs. 0.50 (Low SOB):  $Z=-1.00$ ,  $p=0.317$ . Similarly, in treatment (ii), where SOBs are relatively low, no GA support is found for both men and women: 0.24 (High SOB) vs. 0.20 (Low SOB):  $Z=0.85$ ,  $p=0.396$  and 0.26 (High SOB) vs. 0.27 (Low SOB):  $Z=-0.65$ ,  $p=0.513$ .

## 5. Conclusion

Our study contributes to a large literature in experimental economics that explores gender differences in how people are motivated.<sup>11</sup> With the exception of the study by Kleinknecht (2019), discussed above, no such study that we know of has focused on GA. Our results support the hypothesis that men are more guilt averse than women. The results also support the idea (with which we opened our paper) that a pregnancy-related biological asymmetry could be the underlying cause: In order to get to produce offspring, men need to rely on the trust of women, and GA helps make men trustworthy.

Let us critically assess the relevance of that story. A reason to be skeptical is that there may be *other* situations that favor female GA. Leda Cosmides suggested to us that a man who is investing in a woman’s offspring (post-birth) has to trust that she has been sexually faithful to him. In this situation she may benefit more than he does from being guilt averse. Hence, if there are gender differences, those might be domain-specific.<sup>12</sup>

Our experiment, involving neither pregnancies nor infidelities, does not decouple any stories. We have not explored the relevance of domain-specificity. However, it is possible that evolution, in shaping GA, does not distinguish various trust scenarios all that much. After all, we have found evidence for GA

---

<sup>10</sup> Again, remember that this comparison is not exactly fair, as dictators differ along two dimensions: expectations and the message read.

<sup>11</sup> For pioneering efforts, see Bolton and Katok (1995) and Eckel and Grossman (1998) studies of generosity. Croson and Gneezy (2009) survey 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.

<sup>12</sup> Other couples-stories again favor male guilt. See Dufwenberg’s (2002) analysis of a “marital investment game” featuring a trusting wife with a guilt averse husband for reasons other than pregnancy: 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 towards the end.”

(among men) in our lab setting, which clearly has no exact counterpart in our evolutionarily important distant past. Perhaps differences in GA, for a broad class of situations, rather reflect the frequency with which such a trait is evolutionary advantageous. Our experiment, based on a rather abstract design, may then record decision makers tendencies more broadly. Perhaps our trust-during-pregnancy scenario was relatively frequent back whenever evolution shaped our guilt-tendencies. Perhaps the infidelity story, favoring female guilt, was relevant less often.<sup>13</sup>

Tooby and Cosmides (2008, p. 117) write that “an emotion is a bet placed under conditions of uncertainty: It is the evolved mind’s bet about what internal deployment is likely to lead to the best average long-term set of payoffs, given the structure and statistical contingencies present in the ancestral world”. If our abstract lab design carries external validity to many other naturally occurring settings, then, as regards to GA, evolution seems to have placed its bet on the men.

## References

- Aguiar, F., P. Brañas-Garza, R. Cobo-Reyes, N. Jimenez, and L. M. Miller (2009), “Are women expected to be more generous?” *Experimental Economics*, 12: 93-98.
- Attanasi, G., C. Rimbaud, and M. C. Villeval (2019), “Embezzlement and guilt aversion,” *Journal of Economic Behavior & Organization* 167, 409-429.
- Battigalli, P., and M. Dufwenberg (2007), “Guilt in games,” *American Economic Review*, 97: 170–176.
- Battigalli, P. and M. Dufwenberg (2021), “Belief-dependent motivations and psychological game theory,” *Journal of Economic Literature*, forthcoming.
- Bellemare, C., A. Sebald, and S. Suetens (2017), “A note on testing guilt aversion,” *Games and Economic Behavior*, 102: 233-239.
- Bolton, G. E., and E. Katok (1995), “An experimental test for gender differences in beneficent behavior,” *Economics Letters*, 48: 287–292.
- Borenstein, S. and P. N. Courant (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. and M. Dufwenberg (2006), “Promises and partnership,” *Econometrica*, 74: 1579–1601.
- Childs J. (2012), “Gender differences in lying,” *Economics Letters*, 2012, 114: 147–149.
- 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.

---

<sup>13</sup> Alternatively, other emotions may have evolved to make guilt less relevant in the infidelity scenario. For example, male anger, following detected infidelity, may also deter infidelity. The evolutionary significance of female guilt proneness would thereby be reduced.

- Croson R., and U. Gneezy (2009), “Gender differences in preferences,” *Journal of Economic Literature*, 47: 448–474.
- Dhami S., M. Wei and 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., M. Dufwenberg, and S. Papa (2019a), “The sound of silence: A license to be selfish,” *Economics Letters*, 182: 68–70.
- Di Bartolomeo, G., M. Dufwenberg, S. Papa, and F. Passarelli (2019b), “Promises, expectations & causation,” *Games and Economic Behavior*, 113: 137–146.
- Di Bartolomeo, G., M. Dufwenberg, and S. Papa (2022), “Promises and partner-switch,” mimeo.
- Dreber, A. and M. Johannesson (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. and U. Gneezy (2000), “Measuring beliefs in an experimental lost wallet game,” *Games & Economic Behavior* 30, 163-82.
- Eckel, C. C., and P. J. Grossman (1998), “Are women less selfish than men? Evidence from dictator experiments,” *Economic Journal*, 108: 726–735.
- Ederer, F. and A. Stremitzer (2017), “Promises and expectations,” *Games and Economic Behavior*, 106: 161-178.
- Ellingsen, T., M., Johannesson, S., Tjøtta, and G. Torsvik (2010), “Testing guilt aversion,” *Games and Economic Behavior*, 68: 95–107.
- Khalmetski, K., A. Ockenfels, and P. Werner (2015), “Surprising gifts: Theory and laboratory evidence,” *Journal of Economic Theory*, 159: 163–208.
- Kleinknecht, J. (2019), “A man of his word? An experiment on gender differences in promise keeping,” *Journal of Economic Behavior & Organization*, 168: 251–268.
- Polsby, D.D. and M. Zelder (1994), “Risk-adjusted valuation of professional degrees in divorce”, *Journal of Legal Studies*, 33, 273–285.
- Tooby, J. and L. Cosmides (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*, Feb., 67–68.