The limits of guilt

Loukas Balafoutas, Helena Fornwagner

Working Papers in Economics and Statistics

2016-09

The limits of guilt[#]

Loukas Balafoutas[*]

Department of Public Finance, University of Innsbruck

Helena Fornwagner

Department of Public Finance, University of Innsbruck

Abstract

According to the theory of guilt aversion, agents suffer a psychological cost whenever they fall short of other people's expectations. In this paper we suggest that there may be limits to this kind of motivation. We present evidence from an experimental dictator game showing that behavior is consistent with guilt aversion for relatively low levels of recipient expectations, roughly up to the point where the recipient expects half of the available surplus. Beyond that point the relationship between expectations and transfers becomes negative. We argue that this non-monotonicity can help explain why the economic literature on guilt aversion offers conflicting findings on the relationship between expectations and behavior. Moreover, we examine this relationship at the individual level and establish a typology of subjects depending on how and whether they condition their behavior on recipient expectations. Our evidence is consistent with a simple theoretical model of guilt aversion.

JEL Codes: C91, D03

Keywords: guilt aversion; experiment; strategy method; greed

This version: 14 November 2016

[*] Corresponding author: Department of Public Finance, University of Innsbruck, Universitaetsstrasse 15, A-6020 Innsbruck, Austria. Tel.: +4351250770206. e-mail: loukas.balafoutas@uibk.ac.at

1.   Introduction

Human interaction – in families, companies, or clubs – is often influenced by one's perception of other individuals' expectations. It seems that humans have a tendency to feel guilty when they are letting others down, i.e., when their actions do not meet what they believe others expect from them. This human trait has been coined guilt aversion, defined as the emotion that arises when a player '*believes he hurts others relative to what they believe they will get*' (Charness and Dufwenberg, 2006: 1583).[1] Guilt aversion may influence human behavior in a variety of contexts ranging from marital investments and divorce (Dufwenberg, 2002) to corruption in public administration (Huang and Wu, 1994; Balafoutas, 2011). In an organizational context, relationships between employers and employees can be shaped by mutual expectations on what constitutes appropriate behavior of either party. If both parties are guilt averse, they will consider in their actions (such as choice of wages and work effort) the other party's expectations, thus paving the way for mutually beneficial gift-exchange relationships (Fehr and Gächter, 2000).

In the economic literature guilt aversion is modeled within the analytical framework of psychological game theory (Geanakoplos et al., 1989; Battigalli and Dufwenberg, 2009). In psychological games, players' payoffs depend not only on actions but also on players' (higher-order) beliefs. Consider the simple example where some agent A is asked to make a decision that affects another agent B's payoff (and potentially also A's own payoff): guilt aversion implies that A is more likely to make a decision that increases B's payoff if he believes that B strongly expects him to do so, because A is averse to letting B down and to falling short of her expectations.

Our paper uses a simple design, employing a strategy method variant of the dictator game in order to make two novel contributions to the literature on guilt aversion and more generally on the way that social behavior is affected by (perceived) expectations of the involved parties. First, to the best of our knowledge this is the first study to explicitly put forward the idea that the relationship between expectations and behavior is not necessarily monotonic, but instead has an inverted-U shape on aggregate – as well as at the individual level for some decision makers – in the non-strategic context of the dictator game. We begin by formulating a simple model which predicts this kind of relationship. We then show that dictators display behavior consistent with guilt aversion for relatively low levels of recipient expectations, roughly up to the point where the recipient expects half of the available surplus. Beyond that point, however, the relationship between expectations and transfers becomes negative.

We argue that this non-monotonicity can help explain why the economic literature on guilt aversion, outlined in what follows, offers conflicting findings on the relationship between expectations and behavior. Indeed, if we were to consider only relatively low levels of expectations we would find a significant positive correlation between expectations and transfers consistent with guilt aversion, but this is no longer the case when we use the entire range of expectations. This has led us to talk about 'the limits of guilt': the title of this paper aims to convey the intuitive idea that guilt aversion appears to motivate decision makers, but only up to a certain level. When dictators perceive expectations as being too high and therefore illegitimate, they will not attempt to live up to

---

[1] Similar definitions can be found in the psychology literature, for instance in Baumeister et al. (1995: 173): 'Feeling *guilty [is] associated with...recognizing how a relationship partner's standards and expectations differ from one's own*'.

2

them any longer and they tend to punish recipients who are 'asking too much'. Our work formalizes this idea and presents evidence to support it.

The above intuition is in line with insights from the existing literature on pro-social behavior, for instance with Charness and Rabin (2005) who argue that the extent to which a decision maker responds to the expressed preferences of others depends on how these others have behaved in the past. In particular, they claim that individuals '*are not bothered by disappointing selfish people*' (Charness and Rabin, 2005: 153). Similarly, Ghidoni and Ploner (2014) discuss the idea that only legitimate expectations are worth taking into account by a decision maker. The data presented by Andreoni and Rao (2011) reveal that asking for very high amounts can be counter-productive, in a setting in which recipients can communicate with dictators and make explicit demands. Finally, Regner and Harth (2014) also find an inverse-U shaped relationship between second-order beliefs and the amount returned in a trust game.

Second, our paper establishes a typology of subjects based on examination of the relationship between expectations and behavior at the individual level in the context of the dictator game.[2] It would be unreasonable to suggest that every individual's behavior follows the inverted-U shape described above. Accordingly, we classify the 108 dictators who participated in our experiment into six distinct types: *selfish* types who consistently transfer zero to the recipient; *unconditional altruists* who give a constant positive amount; *positive* (or *guilt averse*) types whose transfers increase with recipient expectations; *negative* types whose transfers decrease with recipient expectations*; hump-shaped* types whose transfers increase with expectations up to a certain (individual-specific) level of expectations and decrease beyond that level, meaning that those subjects display the inverted-U shape also at the individual level; and *other* types who do not fall into any of the five already described categories. We show that positive and negative (monotonic) types account for 18% and 20% of subjects, respectively, while a further 20% are classified as hump-shaped.

Experimental evidence on the role of guilt aversion in decision making has been mixed so far. A number of studies find evidence in favor of guilt aversion (e.g., Dufwenberg and Gneezy, 2000; Charness and Dufwenberg, 2006; Bacharach et al., 2007; Reuben et al., 2009; Dufwenberg et al., 2011; Beck et al., 2013), while others refute it (Vanberg, 2008; Ellingsen et al., 2010 – henceforth EJTT; Kawagoe and Narita, 2014) or find only weak evidence to support it (Charness and Dufwenberg, 2010). A crucial methodological issue concerns the way in which beliefs are measured. Guilt aversion means that a decision maker (DM) suffers a psychological cost whenever she believes that she is falling short of the expectations of an affected party (AP). But how should those second-order beliefs (what the DM thinks the AP expects of her) be measured in an experiment? The approach taken by Charness and Dufwenberg (2006) and others is to elicit the AP's first order belief (what the AP believes the DM will do) and then ask the DM to estimate this first-order belief. This seems like a natural way to elicit second-order beliefs, but a methodological weakness is that it is vulnerable to the so-called false consensus effect. The false consensus effect implies reverse causality, in other

---

[2] Attanasi et al. (2013) provide a categorization of second-movers in a trust game based on the relationship between perceived beliefs and back-transfers. They find that for the majority of subjects back-transfers are conditional on beliefs, and in particular around 55% of subjects are classified as guilt averse. This is an interesting endeavor and it is related to our work, although the trust game features additional possible motivations (compared to the simpler interaction protocol of the dictator game) including strategic behavior and reciprocity.

words that a DM's intended behavior drives her beliefs about what the AP expects through projection.[3] In order to overcome this problem, EJTT do not elicit the DM's second-order beliefs. Instead, they only elicit the AP's first-order beliefs about the DM's giving in the dictator game and then directly transmit those beliefs to the DM. While this procedure does not suffer from a false consensus effect given that beliefs are exogenous to the DM, it opens the door to different kinds of problems: it is conceivable that (some of) the affected parties report beliefs in a strategic manner, for instance if they believe that guilt averse decision makers would then make higher transfers. In this design the AP is namely not informed that his beliefs will be transmitted to the DM, but he is not told the opposite either. Moreover, dictators know that there are undisclosed design features, which may raise suspicion and result in loss of control.[4] In this paper we follow the approach of EJTT, acknowledging however that both methods have their strengths and weaknesses.

We use a methodology very similar to EJTT based on a dictator game, but instead of disclosing first-order beliefs of recipients to dictators and then asking the dictators to decide on a transfer, we ask dictators to report a transfer for each possible first-order belief of the recipient that she is matched with. This technique is akin to a strategy method, since it conditions choices on a co-player's beliefs. Its main advantage is that it allows us to exclude the possibility of a false consensus effect and at the same time to elicit a profile of transfers from the dictator. We are aware of one recent study by Khalmetski et al. (2015) – henceforth KOW – that uses a design very similar to ours, since it employs a strategy method variant of EJTT. This paper is very closely related to our work, especially since the authors find that the relationship between dictator giving and recipient expectations is positive for some dictators and negative for others. What is entirely different, however, is the interpretation of the data. KOW develop a new model of what they call positive and negative surprises. In their framework dictators may have a disutility from creating negative surprises, which leads to a positive relationship between expectations and transfers in line with guilt aversion. But they also enjoy a positive utility from creating positive surprises: the less recipients expect, the greater the positive surprise dictators can create and hence the more they are inclined to give. Notice that this latter motive can lead to a *negative* relationship between expectations and transfers, in line with our results. While we consider this a plausible and interesting story, we note that it is inconsistent with a hump-shaped relationship at the individual level and hence cannot explain the behavior of a substantial fraction of dictators in our sample.[5] Moreover, we go one step further and analyze the relationship between transfers and beliefs at the individual level with the aim of classifying dictators into different types depending on their underlying motivation. Hence, we view our results as complementary to KOW.[6]

---

[3] It is also possible that the relationship between beliefs and actions is contaminated by a self-serving bias. Subjects may also not report truthfully due to risk aversion or strategic considerations. For a detailed discussion on some methodological aspects of belief elicitation see Costa-Gomes and Weizsäcker (2008).

[4] See Khalmetski et al. (2015) for a thorough discussion on this point and also for a robustness check showing that subject behavior remains qualitatively the same when recipients can choose whether to reveal their guesses to the dictators.

[5] See Proposition 1 in KOW: depending on how a dictator weighs positive versus negative surprises in his utility, the relationship between expectations and beliefs can be positive or negative, but it is monotonic in any case.

[6] Hauge (2016) is another recent paper that employs a strategy method variant of the EJTT experiment in which dictators choose their transfers conditional on three possible belief levels (0% of the surplus, 50% of the surplus, or a level in-between). A comparison of our findings to Hauge (2016) is interesting: she finds a positive relationship between transfers and beliefs, which is fully consistent with our findings up to a belief of 8 (50% of the surplus). However, she does not consider higher levels of beliefs, for which we find a negative relationship.

2.   Guilt and its limits in the dictator game

In order to better illustrate the issues discussed in the introduction, we present here a simple model of guilt in a two-person dictator game like the one we use in our experiment. Following Battigalli and Dufwenberg (2007), henceforth B&D, the utility of a dictator i can be defined in terms of a material payoff – which is equal to her endowment e_i minus the transfer t_{ij} made to recipient j – and a disutility from guilt caused by letting recipient j down. How much j is let down is measured by the difference between what j expects to receive and what he actually receives, given by $\alpha_j - t_{ij}$ where α_j denotes the recipient's first-order belief about i's transfer.[7] Following B&D, the parameter $\theta_{ij} \geq 0$ is defined as a measure of i' guilt sensitivity towards j. Then, the utility of dictator i in our experiment would be represented as:

$$u_i = e_i - t_{ij} - \theta_{ij} \max\{0, \alpha_j - t_{ij}\} \qquad (1)$$

We note that, as Battigalli and Dufwenberg (2009) have argued, belief-dependent motivations involving higher-order beliefs can either be modeled using an own belief of a certain order or a belief of another player involving one degree lower order. Accordingly, we use the recipient's first order belief α_j instead of the dictator's second-order belief about α_j. We also note that B&D distinguish between simple guilt and guilt from blame, depending on to the extent to which a decision maker believes that another player believes that he (the decision maker) intended to let her down. In our experimental dictator game the two concepts coincide since the dictator is fully responsible for the payoff allocation and there is no chance move (see Observation 1 in B&D).

Here we modify (1) by introducing a new element into the psychological cost of the dictator. In particular, we posit that this psychological cost depends on a variable H_i, which we define as what i considers to be the *highest legitimate* (or acceptable) *first-order belief* of j. The intuition is simple: each dictator has an individual perception (H_i) of the maximum that her matched recipient is entitled to expect, and the extent to which he suffers guilt from failing to meet the recipient's actual expectation α_j depends on how this actual expectation compares to H_i. The smaller the difference ($H_i - \alpha_j$), the less guilt the dictator suffers from letting the recipient down because in this case α_j is getting closer to the highest legitimate belief; and vice versa for low values of α_j. We use this difference in order to moderate the impact of guilt on i's utility, and define i's utility as follows in our dictator game using a cubic function for the psychological cost arising from guilt:

$$u_i = e_i - t_{ij} - \theta_{ij}\left(H_i - \alpha_j\right)\left(\max\{0, \alpha_j - t_{ij}\}\right)^2 \qquad (2),$$

which can be expressed as

$$u_i = \begin{cases} e_i - t_{ij} - \theta_{ij}(H_i - \alpha_j)(\alpha_j - t_{ij})^2, & \text{if } \alpha_j < H_i \\ e_i - t_{ij}, & \text{if } \alpha_j \geq H_i \end{cases} \qquad (2')$$

---

[7] We follow other applied models of guilt aversion (Beck et al., 2013; Charness and Dufwenberg, 2006, 2011) and take the point expectation α_j as the recipient's reference point against which actual transfers are compared and guilt is evaluated. Hence, the reference point in our specification is not stochastic as, e.g., in KOW.

The intuition behind expression (2) is the following: If $\alpha_j > H_i$, the recipient's belief exceeds the highest legitimate belief and the psychological cost of guilt obtains a positive coefficient, meaning that the dictator would enjoy a positive utility from falling short of the recipient's expectation. In such a case there is no guilt component in the dictator's utility and she would simply transfer zero to the recipient. In other words, the dictator can only suffer guilt when her transfer falls short of beliefs α_j (i.e., when $t_{ij} < \alpha_j$) *and* those beliefs lie below the highest legitimate belief (i.e., when $\alpha_j < H_i$).

Given the above, the optimal transfer for dictator *i* is given by:

$$t_{ij}^* = \max\left\{0, \frac{H_i\alpha_j - \alpha_j^2 - 1/2\theta_{ij}}{H_i - \alpha_j}\right\} \tag{3}$$

Proof: From (2') it follows that $\dfrac{\partial u_i}{\partial t_{ij}} = \begin{cases} -1 + 2\theta_{ij}(H_i - \alpha_j)(\alpha_j - t_{ij}), \text{ if } \alpha_j < H_j \\ -1, \text{ if } \alpha_j \geq H_j \end{cases}$

If $\alpha_j < H_i$ we have a concave parabola in t_{ij} with negative slope at $t_{ij} = \alpha_j$, hence the maximum lies to the left of α_j. Specifically, ignoring constraints $0 \leq t_i \leq e_i$, $\partial u_i / \partial t_{ij}$ yields:

$$t_{ij} = \alpha_j - \frac{1}{2\theta_{ij}(H_i - \alpha_j)} = \frac{H_i\alpha_j - \alpha_j^2 - 1/2\theta_{ij}}{H_i - \alpha_j}$$

Hence, considering that we must have $\alpha_j \leq e_i$ (recipients cannot expect more than the available endowment) and that the unconstrained maximum lies to the left of α_j, we obtain equation (3).

The dictator's optimal transfer as given by (3) is increasing in H_i (the highest legitimate belief) and in θ_{ij} (the guilt sensitivity parameter). More importantly for our research question, the optimal transfer is increasing in α_j but decreasing in α_j^2, which implies an inverted-U shape in the relationship between dictator giving and recipient expectations within the range $0 < \alpha_j < H_i$. Based on equation (3) one can separate between two cases regarding the relationship between dictator transfers and recipient beliefs, depending on the value of H_i:

i. If $H_i > 1/\sqrt{2\theta}$, then the best-reply transfer is first increasing in α_j up to a maximum that depends on H_i (equal to $\hat{\alpha} = H_i - 1/\sqrt{2\theta}$), then it decreases until $\alpha_j = H_i$, and then it is constant at 0 beyond H_i. Hence, for relatively high levels of legitimate expectations, transfers are increasing in beliefs in line with guilt aversion but only up to the point where $\alpha_j = \hat{\alpha}$.

This case predicts the existence of *hump-shaped types* as they are defined in section 4.2: these are dictators who condition their transfers on beliefs in a non-monotonic way following an inverse U. It is worth noting that the model predicts a hump-shaped relationship even if legitimate expectations H_i are very large. As α_j gets closer to H_i from below, the dictator regards α_j as less and less legitimate, and hence he is less willing to give.

6

ii. If, instead, $H_i < 1/\sqrt{2\theta}$, then the best-reply transfer is decreasing in α_j until $\alpha_j = H_i$, and then it is constant at 0 beyond H_i. Hence, for relatively low levels of legitimate expectations, transfers are weakly decreasing in beliefs. This case predicts the existence of *negative types* as they are defined in section 4.2: these are dictators whose transfers are a negative function of recipient expectations.

3. Experimental design and procedures

For our experiment we randomly assigned subjects to one of two types, dictators and recipients, located in two different rooms. The assigned types were fixed, meaning that each subject was either a dictator (called 'Type-A participant' in the instructions) or a recipient (called 'Type-B participant' in the instructions).[8] Dictators received an endowment of €16, while recipients received no endowment (but were paid a show-up fee of €5 following the rules of the lab). Each dictator was then asked to decide how much of their endowment to transfer to the recipient that she had been randomly matched with. Possible transfers included every amount between €0 and 16€ (in €1 steps), including €0 and €16.

Recipients were not able to act at any time during the experiment. However, every recipient was asked about his expectation of the average transfer that dictators would give to recipients within the session. Following EJTT, this was our measure of recipients' first-order beliefs. These estimates were incentivized: The recipient whose expectation was closest to the actual average transfer in the session received €12 in addition to his realized transfer.[9,10] If there was more than one correct estimate, the winner was chosen by chance.

We employed a design akin to the strategy method for dictator decisions. In particular, dictators had to fill out a table where they stated for every possible expectation (i.e., for each elicited first-order belief) of their recipient (varying from €0 to €16) which level of transfer they would like to give. This methodology allows us to elicit a full profile of transfers from each dictator, and for each level of beliefs. Moreover, dictators were given the option to provide us with comments explaining their decisions.[11] Dictators were informed after filling out the table what the estimate of their matched recipient was, and depending on this estimate, the relevant transfer was actually implemented.[12]

---

[8] The full instructions are shown in Appendix D.

[9] For Session 5 we adjusted the lottery to €10.9 (=12€/22*20), because of a lower number of subjects in that specific session (20 instead of 22 dictators and recipients).

[10] Introducing a payment for correct estimates could lead to a bias if subjects start to hedge their experimental income using their stated estimate (Blanco et al., 2008). However, as EJJT note, subjects state their belief about the average realized transfer, and the stakes are small. Therefore the probability of hedging incomes is mitigated. Further, EJTT explain that hedging would only become a problem if the dictators believe that recipients hedge instead of stating their true belief.

[11] Providing comments was however not compulsory, since we did not want to influence subjects by giving the impression that they had to justify their choices to us.

[12] While we are aware of the argument that the strategy method might be prone to demand effects, we note that our results are – to the extent comparable – fully consistent with EJTT's results, where the direct response method is used. Furthermore, numerous studies like Brandts and Charness (2011), Fischbacher et al. (2012), and KOW find no evidence that the two methods yield qualitatively different results.

After having made their choices, subjects of both types were asked to fill out a questionnaire including questions about their person (age, gender, number of siblings). We also asked them some questions regarding the decision situation – for example what they thought the transfer of a dictator should be or, if they were recipients, what they would transfer if they had the role of a dictator. In addition, subjects took a ten-question version of the Big-5 personality questionnaire (Gosling et al., 2003), which analyses personality along five fundamental traits termed extraversion, agreeableness, conscientiousness, neuroticism, and openness. Payments were made anonymously and in cash after filling out all questions. Payments were on average €12.50 for dictators and €9.06 for recipients.

All sessions were conducted at the EconLab of the University of Innsbruck using paper and pen and lasted for around 40 min. We recruited 216 students of different academic backgrounds using H-Root (Bock et al., 2014). We ran five sessions in total, four of them with 44 subjects and one with 40 subjects. This means that we have data for 108 dictators and 108 recipients in total.


4. Results

4.1. Aggregate analysis

Overall, the mean conditional transfer from all 108 dictators in our sample is €3.23, which amounts to 20% of the total available of €16. This is very close to the averages reported in EJTT ($3.60; 24% of the endowment) and KOW (€3.25; 23% of the endowment). Figure 1 plots the mean transfer conditional on each level of beliefs (from 0 to 16). This figure reveals that the relationship between beliefs and dictator giving has an inverted-U shape, with transfers roughly increasing up to a belief of 8 (Spearman's ρ=0.13, p<0.01) and then decreasing for the remaining range of beliefs (ρ=-0.06, p=0.06). This pattern is fully in line with the main hypothesis of the paper that we derived in section 2. It follows that, from the point of view of a recipient, the optimal strategy would be to report an intermediate belief: transfers are highest when beliefs are exactly at the equal split of 8 ($t_8 = 3.75$) and lowest when the recipient expects a transfer of 1 ($t_{15} = 2.50$) – see Table 1, which shows the exact mean transfers by belief level.[13]

Figure 1: Mean transfer, by belief

---

[13] Regarding gender differences in transfers, we report that women in the role of dictators transfer significantly more than men on average (3.30 vs. 3.12, p<0.05, Mann-Whitney U test). This is consistent with much of the relevant literature (see Croson and Gneezy, 2009).
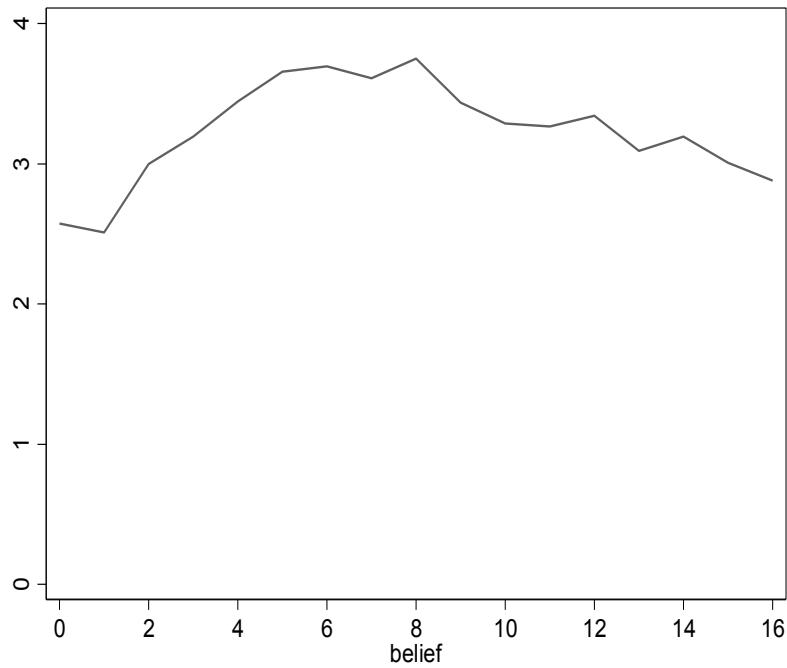
Table 1: Mean transfer, by belief

belief	Mean	Std. Dev.
0	2.57	3.30
1	2.51	2.76
2	3.00	3.04
3	3.19	3.24
4	3.44	3.23
5	3.66	3.33
6	3.69	3.25
7	3.61	3.07
8	3.75	3.42
9	3.44	3.21
10	3.29	3.43
11	3.27	3.44
12	3.34	3.45
13	3.09	3.67
14	3.19	4.07
15	3.01	3.97
16	2.88	4.23
Average	3.23	3.44

The inverted-U shape in the relationship between dictator giving and recipient expectations is the main result of our paper. It is worth pointing out that this result may help explain why a number of papers fail to detect a significant relationship between giving and beliefs, since the increasing and the decreasing part of this relationship are likely to cancel each other out. As a matter of fact, in our experiment we also find no significant correlation between giving and beliefs over the

9

entire range of beliefs (Spearman's ρ=-0.01, p=0.79). Hence, had we only tested for a positive relationship, we would have failed to find one and would have concluded that guilt aversion does not drive dictators' giving decisions.

Table 2 shows the results of Tobit regressions with individual transfers as the independent variable, left-censored at 0.[14] The right-hand side variables are the level of the recipient's belief and its square, in order to control for quadratic effects indicative of an inverted-U shape, as well as age, gender and Big 5 personality traits in specification (2) as control variables. Given that we have 17 observations per subject, all specifications include subject random effects.[15] In both specifications we obtain the predicted positive coefficient for the linear term and negative coefficient for the quadratic term, both significant at the 1% level, confirming our main finding that the relationship between dictator giving and recipient expectations is not monotonic but instead it is positive up to a certain point and then turns negative for high enough expectations. Based on our regression function (1) the global maximum is estimated at a belief level of 7.72, which is in line with the results shown in Table 1. In (2) we include our controls without finding any notable changes in our coefficients of interest.

Table 2: Regression Results

| Dependent variable: dictator giving	(1)	(2)
belief	0.319***	0.329***
	(0.049)	(0.052)
belief^2	-0.020***	-0.021***
	(0.003)	(0.003)
female dictator		-0.536
		(1.041)
age		-0.203
		(0.177)
extraversion		0.028
		(0.371)
agreeableness		0.713
		(0.497)
neuroticism		0.342
		(0.363)
conscientiousness		0.374
		(0.449)
openness		-0.207
		(0.404)
constant	1.414***	0.391
	(0.447)	(6.079)
N	1836	1751

Tobit regressions with dictator random effects. Dependent variable left-censored at 0. Standard errors shown in parentheses. *** denotes significance at the 1% level. As 5 subjects did not fill out the Big Five Questionnaire, the number of observations is lower in specification (2).

---

[14] We have replicated our regressions using OLS and confirmed that they are qualitatively the same.
[15] We are using random effects models instead of fixed effects, since beliefs (and squared beliefs) are constant across dictators. Hence, the assumption of independence between individual heterogeneity and regressors is satisfied and the random effects estimator is efficient.

We conclude this section by showing, in Figure 2, the estimates (first-order beliefs) actually reported by the 108 recipients in our sample regarding the mean transfer from dictators. These estimates vary between €0 and €12 with an average of 4.23 (26% of the endowment). The corresponding mean estimate is €4.70 (34% of the endowment) in KOW, and $4.08 (32%) in EJTT. We also observe gender differences in estimates. In detail, the mean estimate of women is 4.76, which is significantly higher than the mean estimate of 3.82 made by men (p=0.05, Mann Whitney U test). Overall, given that women transfer more as dictators, it is perhaps not surprising that they also report higher estimates for the transfer.

Figure 2: Recipients' beliefs
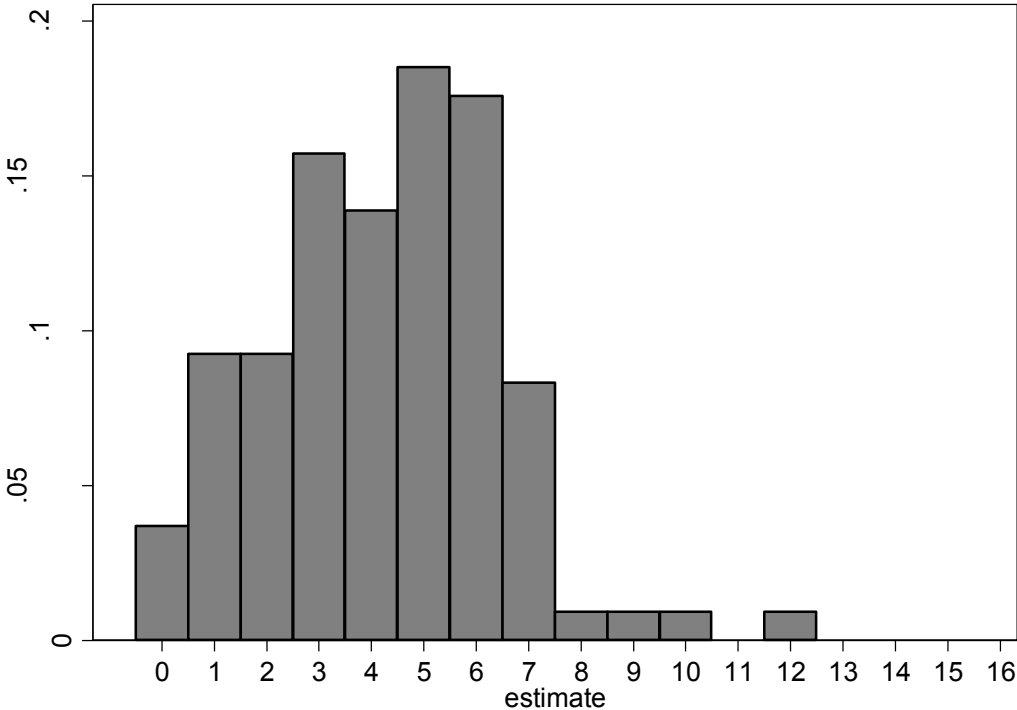
4.2. Individual-level analysis and typology of subjects

In this part we turn to the analysis of the strategy profiles of dictators at the individual level. For this purpose we have plotted the relationship between beliefs and transfers for each dictator and include them in Figure B1 in the Appendix. Based on our model and on the observed patterns of behavior we have classified dictators into one of six distinct behavioral types:

(i) *Selfish* types whose transfers are constant at zero and independent of the recipient's beliefs, with a maximum of one deviation to a positive transfer over the 17 decisions.

(ii) *Unconditional altruists* who transfer a constant positive amount independent of beliefs.

(iii) *Positive (guilt averse)* types whose transfers are positively correlated to recipients' expectations (à la Charness and Dufwenberg, 2006). Following the seminal work by Fischbacher et al. (2001) who classify subjects into four behavioral types based on their strategy profile in a public goods game, we rely on the Spearman rank correlation coefficients and classify a subject as guilt averse if the correlation between transfers and beliefs is positive and significant at least at the 5% level.[16]

(iv) *Negative* types whose transfers are negatively correlated with recipients' expectations (with Spearman's ρ significant at 5%). Remember that, based on our model in section 2, these are dictators with relatively low levels of legitimate beliefs, in particular $H_i > 1/\sqrt{2\theta}$.

(v) *Hump-shaped* types whose transfers are positively correlated with expectations up to a certain threshold, or switching point called S_i, and negatively correlated with expectations beyond S_i (with Spearman's ρ significant at 5% for both). In order to identify these subjects we looked for possible S_i's which would satisfy this condition for each subject, and classified a subject as hump-shaped if such a S_i existed. Based on our model in section 2, these are dictators with relatively high levels of legitimate beliefs, in particular $H_i < 1/\sqrt{2\theta}$.

(vi) *Other* types who do not fall into any of the categories (i) - (v) above.

Hence, two of the above types (selfish subjects and unconditional altruists) do not condition their transfers on the expectations of the recipient, while the opposite is true for types (iii) - (v). Those types condition their transfers on expectations in a systematic way, either positively, negatively, or both.

Table 2 shows the distribution of the six types within the entire population of dictators. The first thing to note is that 20.4% of subjects do not condition their transfers on the expectations of the recipient. Of those, 13.9% are selfish (15 subjects) and 6.5% are unconditional altruists (seven subjects).[17] On the contrary, 58.3% of all subjects conditioned their transfers on expectations in a systematic way. Among those subjects we find a slightly smaller number of guilt averse subjects (with a positive slope in their profile of transfers) than of subjects with a negative slope, with the two types accounting for 17.6% and 20.4% of the sample, respectively. A further 20.4% of subjects can be classified as hump-shaped, i.e., as displaying a positive relationship up to a switching point S_i and a negative one beyond that point. Of course, every one of those dictators may differ with respect to their switching point S_i. In particular, among the 22 subjects in this category, the distribution of the identified levels for S_i is as follows: the mode lies at the equal split of S_i=8 for eight subjects, while two subjects have their switching point at S_i=7 and one subjects at S_i=9, meaning that 50% of subjects belong to that type have their switching point at or around the equal split. Two further subjects

---

[16] Fischbacher et al. (2001) use the 1% significance level as a requirement for their classification. In the Appendix (Table A3) we present a version of Table 2 in which we require that $p<0.01$ instead of $p<0.05$ in order to classify a subject as a positive type, negative type or hump-shaped type. Naturally, this more stringent criterion increases the proportion of subjects who cannot be allocated to one of the five main categories and fall into the category of 'other types'. This affects the classification of 9 subjects in total. Further, two subjects are classified as negative types at the 1% significance level, but became hump-shaped types at the 5% level.

[17] We note that, of the 15 subjects that we classify as selfish, three chose a positive transfer (usually €1) in one of their 17 decisions. We also note that, of the seven subjects that we classify as unconditional altruists, two always chose a transfer of 8 (the equal split) or 1, and the transfer levels of 2, 4 and 6 were each chosen by one subject.

switch already at $S_i=3$, one subject switches at $S_i=4$, three switch $S_i=5$, and five subjects switch at $S_i=6$, respectively.

Table 3: Distribution of types

Person's Type	Freq.	Percent
Selfish	15	13.89
Unconditional altruist	7	6.48
Positive	19	17.59
Negative	22	20.37
Hump-shaped	22	20.37
Other	23	21.30
Total	108	100

4.3. What drives the negative relationship between transfers and beliefs?

The findings discussed so far are in line with our motivation and the hypothesis of our model, which states that the relationship between giving and beliefs is not necessarily monotonic, but includes an upward- and a downward sloping part. While the upward part is consistent with guilt aversion, a motive discussed widely in the literature, the downward sloping part deserves some further deliberation and discussion. We have already indicated in the introduction and in our model that our explanation relies on the idea that recipient expectations can be seen as 'too high' in some cases, leading dictators to reduce their transfers. We have used the word 'greed' to describe this motive and explained that an interesting alternative explanation based on surprises and put forward by KOW cannot account for the hump-shape seen often at the individual level. Here we would like to conclude the results section with a selection of some very characteristic comments made by dictators in our experiments. We offer representative comments related to greed but also comments made by positive (guilt averse) types or by subjects who do not condition their transfer on the recipient's belief, either because they are selfish or because they are unconditional altruists and transfer a constant positive amount – the comment by subject #82 is a good example of such behavior.

Subject # 4 (negative type): '*Player B expects too much.*'

Subject #7 (negative type): '*I will give 0 from a belief of 8 onwards, because if Player B expects that much, she should get nothing*.'

Subject #11 (positive type): '*I make my decision dependent on the estimates of Player B*.'

Subjects #18, #25 (selfish type): '*Sorry I need the money.*'

Subject #45 (hump-shaped): '*Asking 16 is rudeness.*'

Subject #51 (positive type): '*I reward the other if he thinks that I'm generous.*'

Subject # 67 (hump-shaped): '*Other participant demands too much.*'

Subject # 82 (unconditional altruist): 'My decision will not be influenced by the estimate of Player B. B gets an, in my opinion, "fair" amount which should be satisfying.'

Subject # 87 (negative type): '*For beliefs between 11 and 16 my transfer is zero, because these beliefs are too demanding.*'

Subject # 89 (selfish type): '*If I would give something to B, my payment will decrease. My goal is to get as much as possible.*'

Subject # 96 (negative type): '*I do not reward beliefs between 12 and 16.*'

Subject # 100 (positive type): '*Participant B doesn't estimate a high amount, therefore the transfer is low as well. I adjust my transfer for higher estimates of B.*'

Subject # 106 (negative type): '*He expects too much.*'

We believe that the excerpts shown above provide some illustrative evidence in favor of our explanation based on punishing greed for the downwards-sloping part of the inverted-U shape seen in Figure 1. For instance, greed was implicitly (subjects 4, 7, 45, 87, 106) mentioned by some subjects as the reason for reducing transfers when recipient expectations were very high. Given that dictators could provide us with comments regarding their decisions but did not have to do so, only a minority of them took this option (46 subjects in total). Nevertheless, we think it is quite interesting to reproduce here some of those comments (translated from German) along with the type to which each of those subjects is classified. For completeness, in the appendix we provide the full list of all comments made by dictators.

5. Concluding remarks

The goal of this paper has been to make a distinct contribution to the literature on guilt aversion and more generally on psychological games, by suggesting that decision makers who suffer a cost when they fall short of an affected party's expectations will try and live up to these expectations only within certain limits. As we have shown, this implies that the relationship between a decision maker's behavior and an affected party's perceived expectations need not be monotonic. We have used a strategy method variant of the experiment by EJTT and shown that mean transfers across dictators increase with recipient expectations up to a certain threshold but decrease beyond that threshold. This pattern is consistent with a simple theoretical model, which adds the concept of legitimate expectations into the relationship between transfers and beliefs. Furthermore, we have been able to classify dictators into a number of different types depending on the sign of the slope of this relationship in their elicited strategy profile and have found that around six out of ten dictators condition their giving on recipient expectations, either acting in line with guilt aversion, or reducing their transfers as expectations increase, or both.

We believe that, by suggesting that there is a threshold beyond which guilt aversion no longer applies and higher perceived expectations lead to less kind behavior on the part of the decision makers, we are offering an important insight which may help reconcile some of the

14

controversy in the literature on the existence – or not – of guilt aversion. Nevertheless, certain limitations need to be pointed out. For one, we cannot be sure that the mechanism driving the negative part in the relationship between giving and beliefs is due to a motive for punishing greed and recipient expectations that are too high and illegitimate as seen from the perspective of the dictator. We believe that this is a very plausible story and offer some illustrative evidence to support it based on dictators' comments, but readily acknowledge that more evidence is needed in order to corroborate this phenomenon. For instance, one obvious step would be to look for evidence of a role for greed in different contexts, such as the trust games that has been used repeatedly to test for guilt aversion.[18] In any case, we consider our data pattern a very interesting empirical regularity that deserves to be further investigated in future studies.

---

[18] To give one concrete example, we believe that a motive for punishing greed is fully consistent with some of the data patterns presented in the modified trust game in Charness & Dufwenberg (2006). In particular, comparing treatments (5,5) and (7,7) based on game Γ_1 of that paper we see that player B is less trustworthy in treatment (7,7) when the outside options are higher. The authors say that '*perhaps this is...unexpected*' (p. 1588), but we argue that it is reasonable if we consider the idea of legitimate expectations - as defined in our model – from B's point of view. By playing 'In' in (5,5), player A is in effect expecting B to give up 4 so that A can gain 5 (in expected terms). In (7,7) A is in effect asking B to give up 4 so that A can gain only 3, and we conjecture that the lower trustworthiness of player B in this case is because B thinks that A is asking too much.

References

Andreoni, J., Rao, J., 2011. The power of asking: How communication affects selfishness, empathy, and altruism. *Journal of Public Economics* 95, 513-520.

Attanasi, G., Battigalli, P., Nagel, R., 2013. Disclosure of belief-dependent preferences in a trust game. IGIER Working Paper 206, Bocconi University.

Bacharach, M., Guerra, G., Zizzo, D., 2007. The self-fulfilling property of trust: An experimental study. *Theory and Decision* 63, 349-388.

Balafoutas L., 2011. Public beliefs and corruption in a repeated psychological game. *Journal of Economic Behavior and Organization* 78: 51-59.

Battigalli, P., Dufwenberg, M., 2007. Guilt in games. *American Economic Review, Papers and Proceedings* 97(2), 170-176.

Battigalli, P., Dufwenberg, M., 2009. Dynamic psychological games. *Journal of Economic Theory* 114, 1-35.

Baumeister, R., Stillwell, A., Heatherton, T., 1995. Personal narratives about guilt: Role in action control and interpersonal relationships. *Basic and Applied Social Psychology* 17, 173–198.

Beck, A., Kerschbamer, R., Qiu, J., Sutter, M., 2013. Shaping beliefs in experimental markets for expert services: Guilt aversion and the impact of promises and money-burning options. *Games and Economic Behavior* 81, 145-164.

Blanco, M., Engelmann, D., Koch, A.K., Normann, H.T., 2010. Belief elicitation in experiments: Is there a hedging problem? *Experimental Economics* 13, 412-438.

Bock, O., Baetge, I., Nicklisch., A., 2014. hroot – Hamburg registration and organization online tool. *European Economic Review* 71, 117-120.

Brandts, J., Charness, G., 2011. The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics* 14, 375-398.

Charness, G., Dufwenberg, M., 2006. Promises and partnership. *Econometrica* 74, 1579-1601.

Charness, G., Dufwenberg, M., 2010. Bare promises: An experiment. *Economics Letters* 107, 281-283.

Charness, G., Dufwenberg, M., 2011. Participation. *American Economic Review* 101,1211-1237.

Charness, G., Rabin, M., 2005. Expressed preferences and behavior in experimental games. *Games and Economic Behavior* 53, 151-169.

Croson, R., Gneezy, U., 2009. Gender differences in preferences. *Journal of Economic Literature* 47, 448-74.

Costa-Gomes, M., Weizsäcker, G., 2008. Stated beliefs and play in normal-form games. *Review of Economic Studies* 75, 729-762.

Dufwenberg, M., 2002. Marital investments, time consistency and emotions. *Journal of Economic Behavior & Organization* 48, 57-69.

Dufwenberg, M., Gächter, S., Hennig-Schmidt, H., 2011. The framing of games and the psychology of play. *Games and Economic Behavior* 73, 459–478.

Dufwenberg, M., Gneezy, U., 2000. Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior* 30, 163-182.

Ellingsen, T., Johannesson, M., Tjotta, S., Torsvik, G., 2010. Testing guilt aversion. *Games and Economic Behavior* 68, 95-107.

Fehr, E., Gächter, S., 2000. Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives* 14: 159-181.

Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* 71, 397-404.

Fischbacher, U., Gächter, S., Quercia, S., 2012. The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology* 33, 897-913.

Geanakoplos, J., Pearce, D., Stacchetti, E., 1989. Psychological games and sequential rationality. *Games and Economic Behavior* 1, 60-79.

Ghidoni, R., Ploner, M., 2014. When do the expectations of others matter? An experiment on distributional justice and guilt aversion. CEEL Working Paper 3-14.

Gosling, S., Rentfrow, P., Swann, W. Jr., 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality* 37, 504-528.

Hauge, K.E., 2016. Generosity and guilt: The role of beliefs and moral standards of others. *Journal of Economic Psychology*, forthcoming

Huang P., Wu H.M., 1994. More order without more law: A theory of social norms and Organizational cultures. *Journal of Law, Economics and Organization* 10, 390-406.

Kawagoe, T., Narita, Y., 2014. Guilt aversion revisited: An experimental test of a new model. *Journal of Economic Behavior & Organization* 102, 1-9.

Khalmetski, K., Ockenfels, A., Werner, P., 2015. Surprising gifts: Theory and laboratory evidence. *Journal of Economic Theory* 159, 163-208.

Regner, T., Harth, N., 2014. Testing belief-dependent models. Max-Planck Institute of Economics Jena Working Paper.

Reuben, E., Sapienza, P., Zingales, L., 2009. Is mistrust self-fulfilling? *Economics Letters* 104, 89-91.

Vanberg, C., 2008. Why do people keep their promises? An experimental test of two explanations. *Econometrica* 76, 1467-1480.
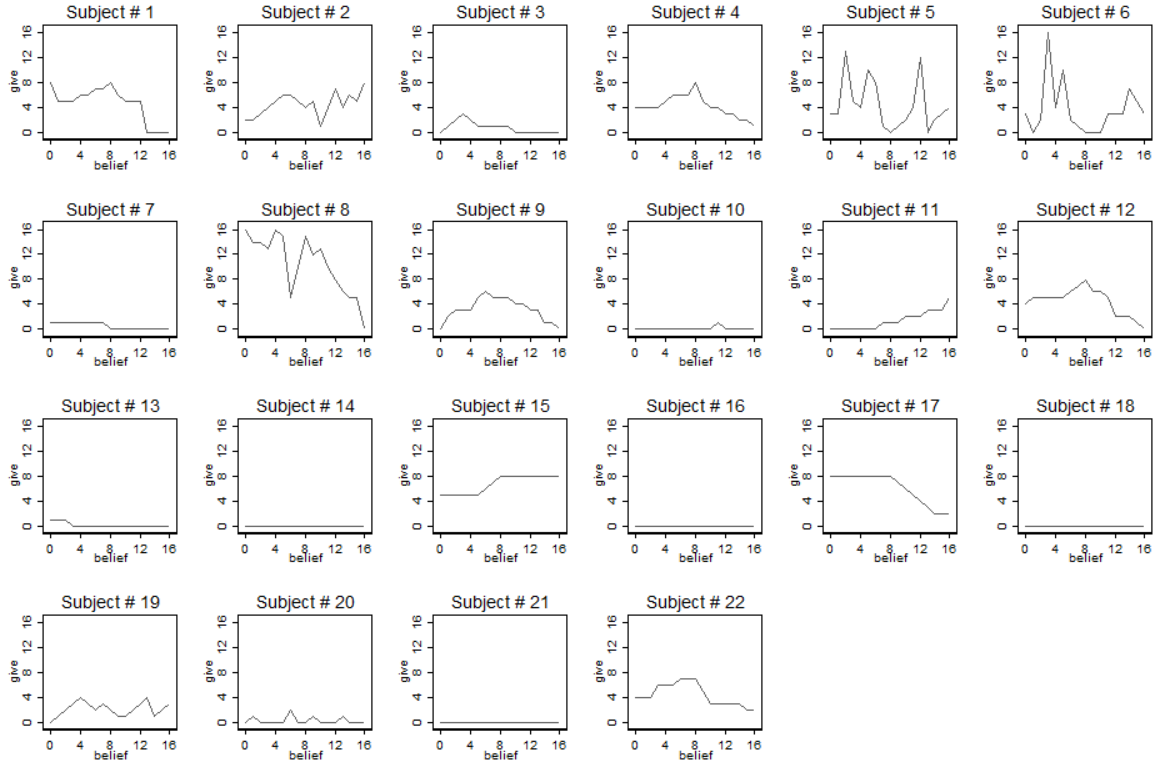
Appendix

A. Tables

Table A1: Distribution of types for different significance levels

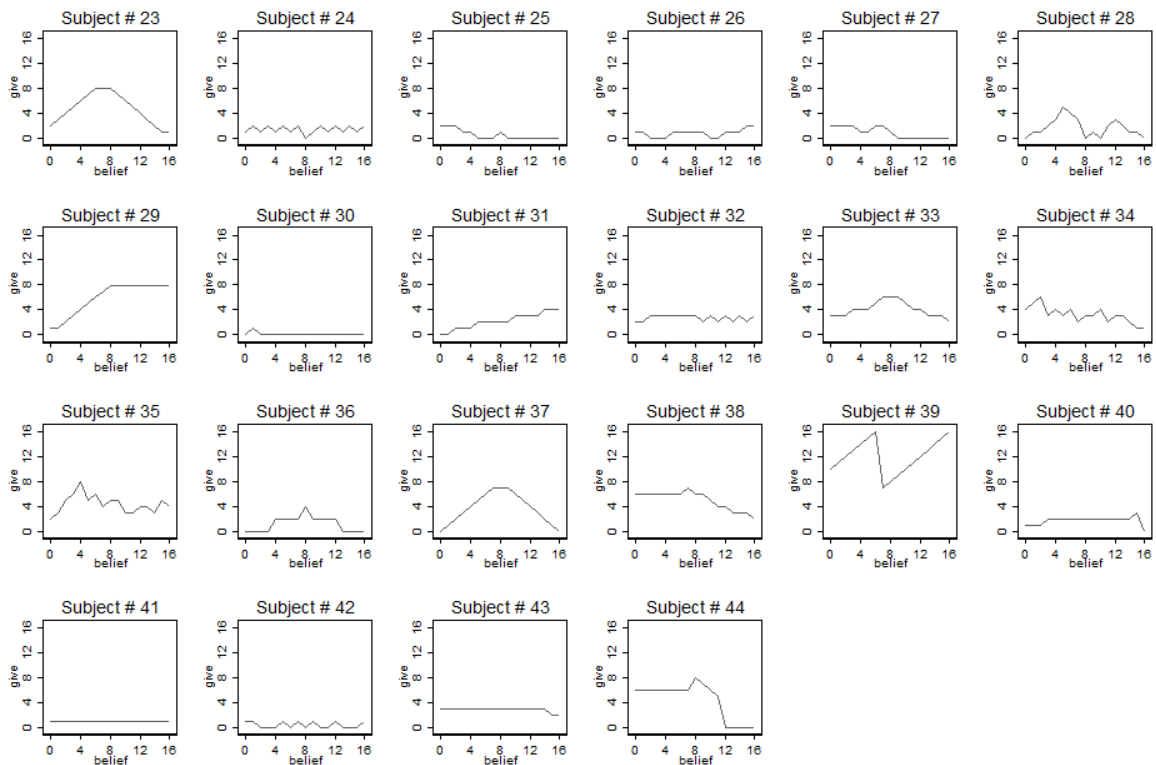|	1% Level		5% Level	
Person's Type	Freq.	Percent	Freq.	Percent
Selfish	15	13.89	15	13.89
Unconditional altruist	7	6.48	7	6.48
Positive	16	14.81	19	17.59
Negative	21	19.44	22	20.37
Hump-shaped	17	15.74	22	20.37
Other	32	29.63	23	21.30
Total	108	100.00	108	100.00

B. Figures

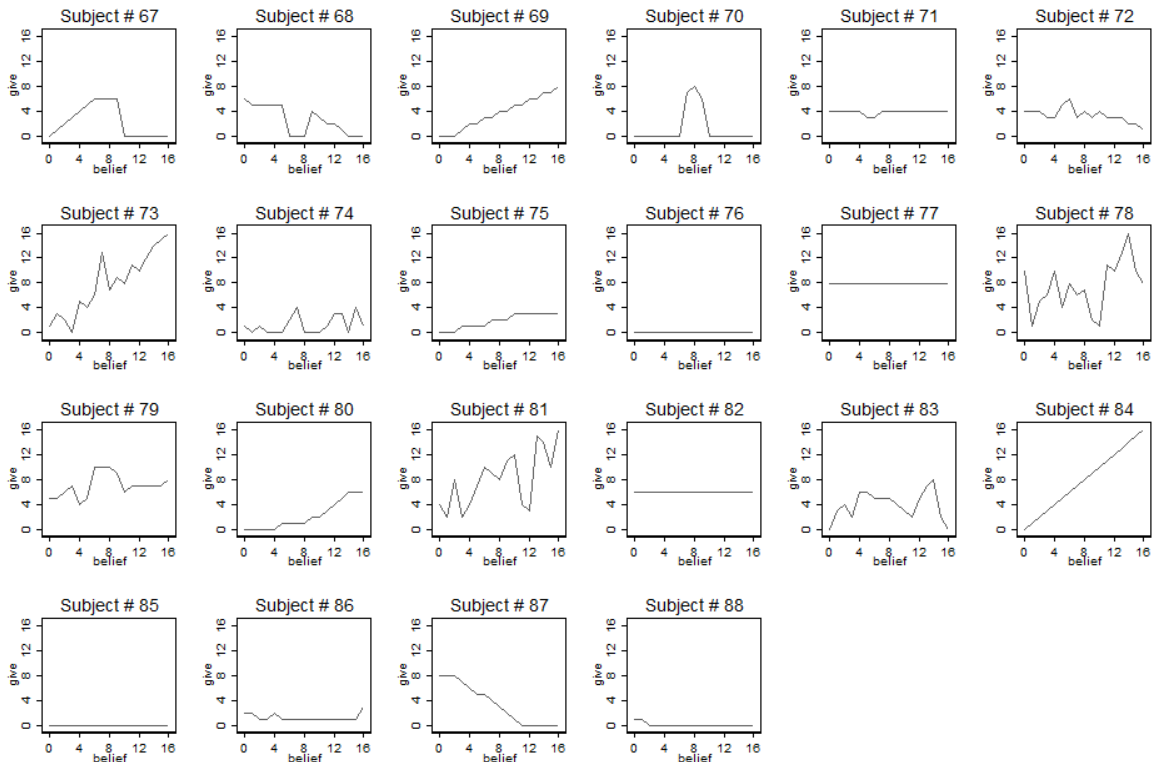**Figure B1: Individual graphs of dictator transfers, by session**

Session 3

Subject # 45, Subject # 46, Subject # 47, Subject # 48, Subject # 49, Subject # 50, Subject # 51, Subject # 52, Subject # 53, Subject # 54, Subject # 55, Subject # 56, Subject # 57, Subject # 58, Subject # 59, Subject # 60, Subject # 61, Subject # 62, Subject # 63, Subject # 64, Subject # 65, Subject # 66

Session 4

Subject # 67, Subject # 68, Subject # 69, Subject # 70, Subject # 71, Subject # 72, Subject # 73, Subject # 74, Subject # 75, Subject # 76, Subject # 77, Subject # 78, Subject # 79, Subject # 80, Subject # 81, Subject # 82, Subject # 83, Subject # 84, Subject # 85, Subject # 86, Subject # 87, Subject # 88
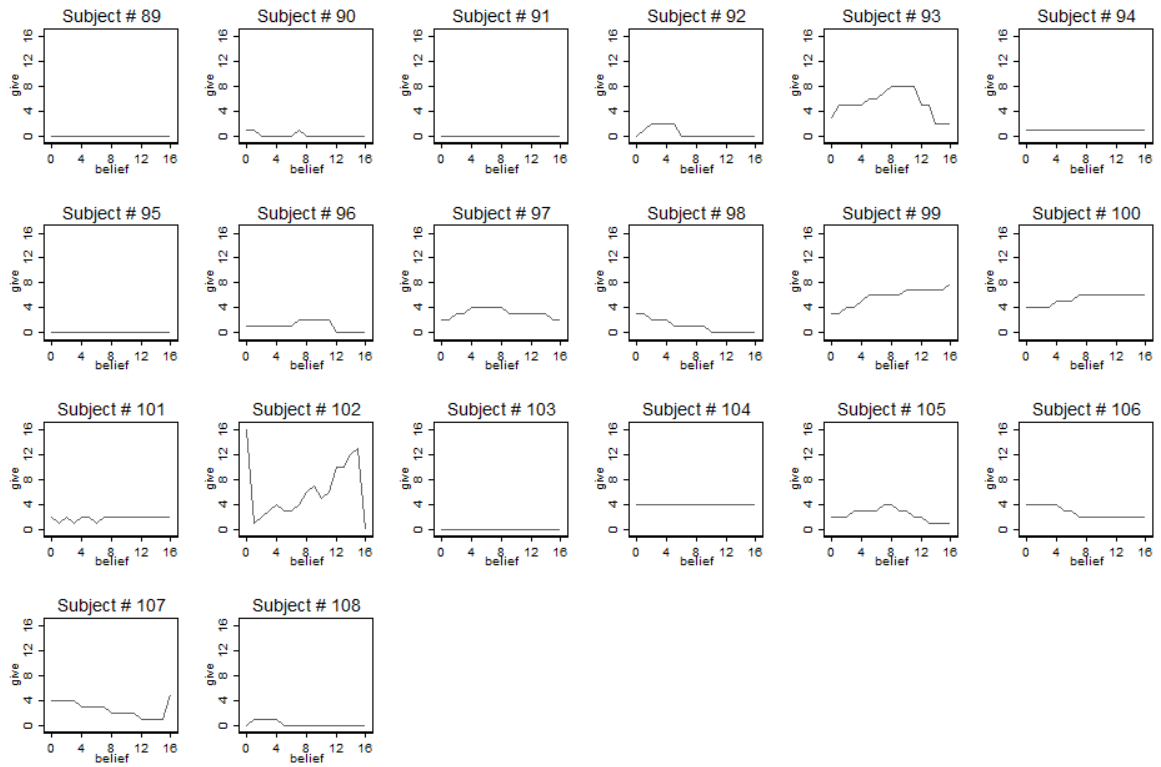
Session 5

C. Full list of comments

Session 1

Subject # 3   (on belief 0-3): Even if the average is not easy to influence, I want be next to the average.

(on belief 10-16): I think this estimates are not realistic.

Subject # 4   (on belief 8): If he expects a fair split, I appreciate this. In this case I want to split equally.

(on belief 16): I think in this case he wants to much.

Subject # 6   (on belief 5): 5 is half of 10 :-)

Subject # 7   (on belief 16): I would give the participant-B 0€ if he would expect a transfer of 16€, because following I would get nothing.

Subject # 9   (on belief 0): Only 0€, because I would feel sorry if he beliefs that I would give him nothing.

(on belief 0-16): In general I my transfers are higher to the middle and are decreasing to the endings.

Subject # 11 (on belief 0-16): My decisions depend on the beliefs of the other.

Subject # 13   (on belief 0-2): Not much, but I think the other write 0€.

(on belief 3-5): I think more than 3/4 expects nobody.

Subject # 18 (on belief 0-16): I need the money.

Subject # 19   (on belief 0): If he expects zero I will give him zero.

(on belief 6): I reward a realistic estimate.

Session 2

Subject # 24   (on belief 0): As a faire reward.

(on belief 8): He probably will estimate this.

(on belief 16): Becomes 2€ if he really thinks I would transfer him 16€.

Subject # 25   (on belief 0-16): Sorry need the money.

Subject # 26   (on belief 0-16): I think the others will do the same.

Subject # 27   (on belief 9-16): I don't think participant B thinks I will transfer him more than half.

Subject # 29   (on belief 0): Something he/she should get.

(on belief 1-7): estimate=transfer from me (he / she will receive as much as he / she estimates --> no frustration

(on belief 8): Everyone of us receives the same amount (this is fair)

(on belief 9-16): Nobody should have more than the other.

Subject # 34   (on belief 16): I don't think somebody expects everything.

Subject # 38   (on belief 0): Nobody is this cruel.

(on belief 7): Realistic estimate.

(on belief 16): Unrealistic estimate.

Subject # 40   (on belief 0-15): The other would estimate 2€ or less because students are all about the money.

(on belief 16): Much to high estimate.

Subject # 41   (on belief 0-16): Sorry I'm bankrupt.

Session 3

Subject # 45   (on belief 3-7): Fair.

(on belief 16): For audacity.

Subject # 51   (on belief 0-16): The idea behind this is to support the trust in human being - if he thinks we are this generous, we really give for example 3€ and so brave is to really estimate this, he should be rewarded - even if I don't get the highest possible outcome or even nothing.

Subject # 52   (on belief 13-16): I don't think B expect his, nevertheless if 0€ or 16€ is possible.

Subject # 56   (on belief 0-16): Participant B will not be more or less sympathetic for me because of his estimate and therefore this will not end up in a change of my transfer. I think 2€ are a good payment for him giving up some leisure time.

Subject # 58   (on belief 0-6): He / she could get 12€.

Subject # 59   (on belief 0): Really pessimistic (or unrealistic)?

(on belief 3-8): I'm young an need the money.

(on belief 9-16): Unrealistic estimate.

Subject # 60   (on belief 0-16): No risk, no fun.

Subject # 62   (on belief 0-16): I don't think he expects 0€.

Subject # 63   (on belief 0-16): I would expect this if I would be participant B; therefore the chance to receive 12€.

Subject # 65   (on belief 0-16): My utility maximization! I'm not interested in sending any positive amount to a stranger if the whole experiment is anonymous.

Session 4

Subject # 67   (on belief 0-6): estimate is right, chance for € 12

(on belief 7-9): I want to have at least € 10.

(on belief 10-16): Other participant demands too much.

Subject # 68   (on belief 0-5): Payment for B without big loss for A

(on belief 6-8): Estimated Average, Winning for B

Subject # 70 (on belief 7-9): best for both would be fair splitting, therefore the other participant gets only something if he acts fair

Subject #71 (on belief 0-16): Otherwise the payment for B wouldn't be fair, at least 25 %

Subject #72 (on belief 0): B shouldn't be here for nothing

(on belief 16): Way to high expectations

Subject # 74 (on belief 0): endowment

Subject #82 (on belief 0-6): My decision will not be influenced by the estimate of the B. B gets an, in my opinion, "fair" amount which should be satisfying.

Subject #85 (on belief 0): Experiment Participant

Subject #87 (on belief 0): € 8 because of humility

(on belief 3-7): I don't want to give more than the half, because I really need the money. But probably I would anyhow give him more than the half if he wants more.

(on belief 11-16): Zero, would be too demanding for me

Session 5

Subject # 89 (on belief 0-16): If would I give something to Player B my payment will decrease. My goal is to get as much as possible.

Subject #93 (on belief 0-5): very pessimistic estimate

(on belief 6-10): Fair estimate

(on belief 14-16): too much

Subject #94 (on belief 0-16): Because the transfer is made anonymously I decided like this. If the other person would know his opponent I would have made a fair transfer.

Subject #96 (on belief 0-6): Arbitrariness, I was often in the B position in this kind of experiment

(on belief 7-11): Self-confidence and courage are being rewarded

(on belief 12-16): No reward for this

Subject #97 (on belief 0-3): more than he expected

(on belief 4-7): I would give him this much

(on belief 8): exactly half of the amount

(on belief 9-16): giving more than to keep for oneself is unrealistic

Subject #98 (on belief 0-9): I want to leave this experiment with as much money as possible, because of this I will give maximum € 3.

(on belief 10-16): I would give € 0 here because I don't think that the estimate will be more than € 10

Subject #100 (on belief 0-3): Participant B doesn't estimate a high amount, therefore the transfer is low as well.

(on belief 4-6): adjusted transfers amount to higher estimates of B

(on belief 7-16): Estimate is rising, Transfer as well, but since it's anonymous and can decide, I want to keep more for myself. And also I think such a high estimate is not suitable and not realistic.

Subject #106 (on belief 0-4): no stingy estimate but I don't want to give more

(on belief 9-16): He expects too much

Subject #108 (on belief 0-5): he would be near to the average estimate

(on belief 6-16): nobody will estimate this much

D. Instructions – *not intended for publication*

a. Dictators

Welcome to our experiment and thank you for your participation. Please do not talk to any other participants during the experiment.

Instructions

During the experiment you and all other participants have to make some decisions. Depending on the decisions you are able to earn money. You will receive the money earned, dependent on your decisions, anonymous and in cash at the end of the experiment.

If you have any question after we read the instruction to you, please raise your hand. The experimenter will come to you and will answer your question individually. The duration of the experiment is calculated with around 30 min.

2 Types of participants

There are two types of participants: Type A and Type B. Every participant in this room was randomly assigned to be a participant of Type A. Every Type A is randomly matched with a Type B participant who is located in the room next door. Your will never get to now with whom you interacted, neither during nor after the experiment. The other participants will never get to know which decisions you took or what you have earned and your identity.

The basic decision

Every Type A participant receives an endowment of €16. Type B participants receive no endowment. Participant A decides for an amount of money of his / her endowment which he / she wants to transfer to the participant B that he / she is matched with. Every amount between €0 and €16, including €0 und €16 can be transferred (in 1 € – steps). This means:

Earnings participant type A = €16 – Transfer to B-participant

Earnings participant type B = Transfer received from participant type A

Participant B cannot act. But every B-participant is asked, before an A-participant makes the decision, about his / her estimate of the average transfer an A-participant gives to a B-participant. The A-participants are informed after their decisions what the estimate of their matched B-participant was, but they can relate their transfer to the different estimates.

In other words:

You are telling us by the use of the attached table which transfer you would like to give your B-participant for each level of his / her estimate.

Depending on the level he / she really estimated the related transfer from the table will be actually transferred.

The B-participants do not know, that the A-participants will be informed about their estimates.

In addition, B-participants can earn further money depending on their estimate: The B-participant whose estimate is closest to the actual average transfer from A-participants will receive another €12 in addition to their realized transfer (if there is more than one correct estimate, the winner will be chosen by chance).

Please write down your seating number: _____

Your B-participant expects the following transfer from you:	What do you really want to transfer to your B-participant? (in integers)	Comments
€0		
€ 1		
€ 2		
€ 3		
€ 4		
€ 5		
€ 6		
€ 7		
€ 8		
€ 9		
€ 10		
€ 11		
€ 12		
€ 13		
€ 14		
€ 15		
€ 16		

Welcome to our experiment and thank you for your participation. Please do not talk to any other participants during the experiment.

Instructions

During the experiment you and all other participants have to make some decisions. You are getting 5€ for your participation, independently of which decisions you take in the following experiment. But in addition, dependently on your decisions you are able to earn money.

You will receive the money earned, dependent on your decisions, anonymously and in cash at the end of the experiment. If you have any question after we read the instructions to you, please raise your hand. The experimenter will come to you and will answer your question individually. The duration of the experiment is calculated with around 30 min.

2 Types of participants

There are two types of participants: Type A and Type B. Every participant in this room was randomly assigned to be a participant of Type B. Every Type B is randomly matched with a Type A participant who is located in the room next door. Your will never get to now with whom you interacted, neither during nor after the experiment. The other participants will never get to know which decisions you took or what you have earned and your identity.

The basic decision

Every Type A participant receives an endowment of €16. Type B participants receive no endowment. Participant A decides for an amount of money of his / her endowment which he / she wants to transfer to the participant B that he / she is matched with. Every amount between €0 and €16, including €0 und €16 can be transferred (in 1 € – steps). This means:

Earnings participant type A = €16 – Type B

Earnings participant type B = Transfer received from participant type A

Participant B cannot act. But every B-participant is asked, before A-participants make their decisions, about his / her estimate of the average transfer that A-participants will give.

In addition, B-participants can earn further money depending on their estimate: The B-participant whose estimate is closest to the actual average transfer from A-participants will receive another €12 in addition to their realized transfer (if there is more than one correct estimate, the winner will be chosen by chance).

In other words:

We are asking you to report on the attached sheet of paper your estimate of the average transfer A-participants make to B-participants.

[Recipients' decision sheet:]

Please write down your seating number   _____

Please decide now on your estimate:

What do you think is the average transfer a participant of type A makes to a participant of type B.

Please state your estimate in full € steps – you can write down every amount from 0€ to 16€, including 0€ and 16€.

Your estimate is: _____

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

2016-09 **Loukas Balafoutas, Helena Fornwagner:** The limits of guilt

2016-08 **Markus Dabernig, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Spatial ensemble post-processing with standardized anomalies

2016-07 **Reto Stauffer, Jakob W. Messner, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Spatio-temporal precipitation climatology over complex terrain using a censored additive regression model

2016-06 **Michael Razen, Jürgen Huber, Michael Kirchler:** Cash inflow and trading horizon in asset markets

2016-05 **Ting Wang, Carolin Strobl, Achim Zeileis, Edgar C. Merkle:** Score-based tests of differential item functioning in the two-parameter model

2016-04 **Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Non-homogeneous boosting for predictor selection in ensemble post-processing

2016-03 **Dietmar Fehr, Matthias Sutter:** Gossip and the efficiency of interactions

2016-02 **Michael Kirchler, Florian Lindner, Utz Weitzel:** Rankings and risk-taking in the finance industry

2016-01 **Sibylle Puntscher, Janette Walde, Gottfried Tappeiner:** Do methodical traps lead to wrong development strategies for welfare? A multilevel approach considering heterogeneity across industrialized and developing countries

2015-16 **Niall Flynn, Christopher Kah, Rudolf Kerschbamer:** Vickrey Auction vs BDM: Difference in bidding behaviour and the impact of other-regarding motives

2015-15 **Christopher Kah, Markus Walzl:** Stochastic stability in a learning dynamic with best response to noisy play

2015-14 **Matthias Siller, Christoph Hauser, Janette Walde, Gottfried Tappeiner:** Measuring regional innovation in one dimension: More lost than gained?

2015-13 **Christoph Hauser, Gottfried Tappeiner, Janette Walde:** The roots of regional trust

2015-12 **Christoph Hauser:** Effects of employee social capital on wage satisfaction, job satisfaction and organizational commitment

2015-11 **Thomas Stöckl:** Dishonest or professional behavior? Can we tell? A comment on: Cohn et al. 2014, Nature 516, 86-89, "Business culture and dishonesty in the banking industry"

2015-10 **Marjolein Fokkema, Niels Smits, Achim Zeileis, Torsten Hothorn, Henk Kelderman:** Detecting treatment-subgroup interactions in clustered data with generalized linear mixed-effects model trees

2015-09 **Martin Halla, Gerald Pruckner, Thomas Schober:** The cost-effectiveness of developmental screenings: Evidence from a nationwide programme

2015-08 **Lorenz B. Fischer, Michael Pfaffermayr:** The more the merrier? Migration and convergence among European regions

2015-07 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city

2015-07 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city *forthcoming in European Economic Review*

2015-06 **Martin Geiger, Wolfgang Luhan, Johann Scharler:** When do Fiscal Consolidations Lead to Consumption Booms? Lessons from a Laboratory Experiment

2015-05 **Alice Sanwald, Engelbert Theurl:** Out-of-pocket payments in the Austrian healthcare system - a distributional analysis

2015-04 **Rudolf Kerschbamer, Matthias Sutter, Uwe Dulleck:** How social preferences shape incentives in (experimental) markets for credence goods *forthcoming in Economic Journal*

2015-03 **Kenneth Harttgen, Stefan Lang, Judith Santer:** Multilevel modelling of child mortality in Africa

2015-02 **Helene Roth, Stefan Lang, Helga Wagner:** Random intercept selection in structured additive regression models

2015-01 **Alice Sanwald, Engelbert Theurl:** Out-of-pocket expenditures for pharmaceuticals: Lessons from the Austrian household budget survey

University of Innsbruck

Working Papers in Economics and Statistics

2016-09

Loukas Balafoutas, Helena Fornwagner

The limits of guilt

Abstract
Guilt aversion has been put forward in recent years as a prominent motivation for certain aspects of human behavior. When agents are guilt averse, their utility depends on what they believe others expect of them and they suffer a cost whenever they fall short of those expectations. In this paper we suggest that there may be limits to this kind of motivation. We present evidence from a dictator game showing that dictators display behavior consistent with guilt aversion for relatively low levels of recipient expectations, roughly up to the point where the recipient expects half of the available surplus. Beyond that point the relationship between expectations and transfers becomes negative. We argue that this non-monotonicity can help explain why the economic literature on guilt aversion offers conflicting findings on the relationship between expectations and behavior. Moreover, we examine this relationship at the individual level and establish a typology of subjects depending on how and whether they condition their behavior on recipient expectations. Our evidence is consistent with a simple theoretical model of guilt aversion.