CESifo WORKING PAPERS

6999
2018

April 2018

Norms and Guilt

Anastasia Danilov, Kiryl Khalmetski, Dirk Sliwka

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6999*
Category 13: Behavioural Economics

Norms and Guilt

Abstract

It has been argued that guilt aversion (the aversion to violate others' expectations) and the compliance to descriptive social norms (the aversion to act differently than others in the same situation) are important drivers of human behavior. We show in a formal model that both motives are empirically indistinguishable when only one benchmark (another person's expectation or a norm) is revealed as each of these benchmarks signals information on the other one. To address this problem, we experimentally study how individuals react when both benchmarks are revealed simultaneously. We find that both types of information affect transfers in the dictator game. At the same time, the effect of the recipient's expectation is non-monotonic as dictators use the disclosed expectation in a self-serving way to decrease transfers.

JEL-Codes: C910, D830, D840.

Keywords: guilt aversion, social norms, conformity, dictator game.

Anastasia Danilov
University of Cologne, Department of Corporate Development and Business Ethics, Albertus-Magnus-Platz
Germany – 50923 Cologne
danilov@wiso.uni-koeln.de

Kiryl Khalmetski
University of Cologne
Department of Economics
Albertus-Magnus-Platz
Germany – 50923 Cologne
kiryl.khalmetski@uni-koeln.de

Dirk Sliwka
University of Cologne
Department of Economics
Albertus-Magnus-Platz
Germany – 50923 Cologne
sliwka@wiso.uni-koeln.de

April 17, 2018
We thank Gary Charness, Erin Krupka and RobertoWeber for helpful comments. Financial support of the German Research Foundation (DFG) through the Research Unit "Design and Behavior" (FOR 1371) is gratefully acknowledged.

1 Introduction

In situations involving social interactions people tend to consider specific benchmarks indicating "appropriate" behavior in a given context. One such benchmark is the perception about the action commonly chosen by other decision makers in the same situation, or the *descriptive social norm* of behavior. Another important benchmark is the action that is expected by another party directly affected by the choice (e.g., a waiter expecting a certain amount as a tip, an employee expecting a specific wage increase). A preference for compliance to such individual expectations is termed as *guilt aversion* (Battigalli and Dufwenberg, 2007).[1]

A large body of literature has acknowledged both of these mechanisms to be important determinants of individual behavior.[2] Yet, there is still a significant gap in understanding how the effects of these two behavioral benchmarks interact, or whether they may even substitute each other. The streams of literature studying the impact of either norm compliance or guilt aversion usually focus on one of the two benchmarks. However, these concepts can be closely interrelated in several ways. First, when only one benchmark is known, individuals can make inferences about the other benchmark and change their beliefs and actions accordingly. For instance, a well-known social norm can indicate what kind of behavior an average person expects to observe in a given situation. Taking an example from Charness and Dufwenberg (2006), a waiter, who works in a place where a tipping norm of 15% applies, might generally expect to be tipped according to this social norm. Conversely, revealed expectations of another party may provide a signal about the prevalent social norm when it is unknown (for instance, in a new cultural environment or a different organization).[3] Hence, the experimental manipulation of just one of these two benchmarks will not be sufficient to separate the direct effect from the indirect effect as one of these benchmarks is also a signal of the other one. In this paper we thus study the interplay between both these benchmarks.

As a first step we formalize the claim that descriptive social norms and individual expectations are mutual signals of each other within a stylized learning model in

[1] The literature typically distinguishes between descriptive (how others mostly behave) and injunctive norms (what others mostly believe is appropriate) (see Cialdini et al., 1990). Since we subsequently study the interaction between norm conformity and guilt aversion, our focus is on the descriptive norms as guilt aversion is mainly studied through subjects' expectations about the actual, and not normative behavior. Hence, the confounding correlation between norms and expectations, as discussed in details below, is more relevant for conformity to descriptive norms. Moreover, descriptive norms have been also proven to be a better predictor of the individual behavior than injunctive norms (Bicchieri and Xiao, 2009).

[2] See Bardsley and Sausgruber (2005), Krupka and Weber (2009), Bicchieri and Xiao (2009) for norm conformity, and Dufwenberg and Gneezy (2000), Charness and Dufwenberg (2006), Khalmetski et al. (2015) for guilt aversion. Note, however, that Ellingsen et al. (2010) do not find evidence in line with guilt aversion.

[3] See Sliwka (2007), Friebel and Schnedler (2011), Van der Weele (2009), and Bénabou and Tirole (2012) for a theoretical analysis and Danilov and Sliwka (2016) for corresponding experimental evidence that revealed choices of informed principals might signal social norms to the agents.

Section 2. In Section 3, we describe an experiment building on the dictator game to disentangle the 'pure' effects of norm compliance and guilt aversion (controlling for mutual signaling). In particular, we study the behavior of dictators when *both* the expectations of the recipient and the descriptive social norm are disclosed. To do this, the dictator game itself is repeated twice with different recipients. Before the second game dictators learn *both,* the recipient's expectations and the average transfer in a group of other players in the same role in the previous game, prior to making their own transfer choice. In this way, we eliminate mutual signaling effects and, hence, can disentangle the effects of norm conformity and guilt aversion from each other. Besides, our design elicits expectations in an incentive compatible manner, avoids deception, and allows us to control for subjects' own social preferences.

We find that (i) descriptive norms have a substantial effect on dictators' giving rates and (ii) the recipient's expectation affects dictators' behavior, yet in a non-monotonic way such that too high expectations even lead to a reduction of transfers. The results are thus well in line with the coexistence of norm compliance and guilt aversion. However, we also find that recipient's expectations are used in a rather self-serving manner. Low expectations of a recipient seem to be used as an excuse to lower transfers and high expectations are rather punished. In contrast, descriptive norms monotonically increase transfers.

To the best of our knowledge, our study is the first to study simultaneously the causal effects of information about (i) the actual behavior of others in the same role, as well as (ii) the expectations of a directly affected other player to disentangle compliance to descriptive norms from guilt aversion. It is related to several recent papers on the role of norms and guilt aversion. Krupka et al. (2017) study how ex-ante informal agreements affect behavior in symmetric games. They argue that such agreements may affect behavior through changes in *injunctive norms* (i.e., the commonly shared beliefs about the socially appropriate behavior) as well as by changing the expectations of the respective other party. The authors do not exogenously vary information about norms and beliefs but estimate structural models using information about injunctive norms elicited in a second experiment. They show that predictive power is highest when both motives are taken into account. Hauge (2016) studies the effects of injunctive norms and recipients' expectations in the dictator game separately without comparing their effects in the same treatment.

Related to our findings on the non-monotonic effect of induced second-order beliefs, some recent studies also show that the strength of guilt aversion depends on the "reasonableness" of expectations. Pelligra et al. (2016) provide experimental evidence that the returned amounts of trustees in a trust game are not affected by reported expectations of the experimenter if these expectations exceed trustees' prior (i.e. unconditional) return amounts. Khalmetski (2016), based on his experimental data, suggests that subjects might tend to ignore higher expectations of another player if the latter are based on a different information than the one the decision maker has. Balafoutas and Fornwagner (2017) study the effect of exogenously varied recipient expectations in a dictator game. They find that the effect of expectations on giving gets even negative for a substantial share of dictators if these expectations exceed a certain level (which in turn varies across subjects).

The rest of the paper is organized as follows. Section 2 provides a model formalizing mutual signaling of norms and expectations, Section 3 describes the experimental design and hypotheses, Section 4 presents the results, while Section 5 provides their discussion and concludes.

2 Model

2.1 Baseline setting

Consider an agent i who takes a personally costly action $a_i \in [0,1]$ which increases the utility of another agent j. The agent is potentially guilt averse and may have a preference for norm compliance such that her utility function is

$$u\left(a_i, E_j, N, \theta_{Gi}, \theta_{Ni}\right)$$

where $E_j = E_j\left[a_i\right]$ is the expectation of agent j about i's action, and N is a descriptive social norm. The social norm N is equal to the population mean of the chosen action $E\left[a\right]$, which is unknown to the agent. The agent's type is determined by $\theta_{Gi}, \theta_{Ni} \in \mathbb{R}_0^+$, where θ_{Gi} measures the degree of guilt aversion and θ_{Ni} the propensity for norm compliance.

We furthermore assume that

- $\frac{\partial u}{\partial a_i} < 0$ for $\theta_{G_i}, \theta_{N_i} = 0$,
- $\frac{\partial u}{\partial a_i} > 0$ for $a_i < E_j$ if θ_{Gi} is larger than some cutoff $\bar{\theta}_G$,
- $\frac{\partial u}{\partial a_i} > 0$ for $a_i < N$ if θ_{Ni} is larger than some cutoff $\bar{\theta}_N$,
- $\frac{\partial u}{\partial a_i} < 0$ for $a_i > \max\left\{E_j, N\right\}$.

An example for such a utility function is

$$u\left(a_i, E_j, N, \theta_{Gi}, \theta_{Ni}\right) = K - a_i - \theta_{Gi} \max\left\{0, E_j - a_i\right\} - \theta_{Ni} \max\left\{0, N - a_i\right\}$$

where K is some parameter, which nests a standard linear guilt aversion model (if $\theta_{Ni} = 0$). But the assumptions also allow for non-linear psychological costs of deviation from expectations or norm violation.

We now show that when there is uncertainty about either j's expectation E_j or the norm N the following holds, respectively:

(i) When an agent i learns a signal about j's expectation but is *not* guilt averse (i.e. $\theta_{Gi} = 0$), then a_i is still increasing in this signal if i is sufficiently norm compliant.

(ii) When an agent i learns a signal about the norm but is *not* norm compliant (i.e. $\theta_{Ni} = 0$), then a_i is still increasing in this signal if i is sufficiently guilt averse.

In other words, norm compliance generates guilt averse behavior when there is uncertainty about the norm. And vice versa, guilt aversion produces norm-compliant behavior when there is uncertainty about the other player's expectation. The argument is formalized in the subsequent sections.

2.2 Information structure

We consider the following information structure. Assume that agents ex-ante believe that N is distributed on $[0,1]$ according to a cumulative distribution function $G(N)$ with probability density $g(N)$.[4] Every agent i gets a noisy signal $s_i \in S \subset \mathbb{R}$ about the norm, which is independently and identically distributed according to a cumulative distribution function $F(s|N)$ with probability density $f(s|N)$. These signals can be interpreted, for instance, as some prior private knowledge about the strategic setting. All the functions are assumed to be continuously differentiable in all arguments. Besides, we assume that $F(s|N)$ satisfies the strict monotone likelihood ratio property (MLRP; Milgrom, 1981), i.e.,

$$\forall s \in S, N'' > N' : \frac{\partial}{\partial s} \frac{f(s|N'')}{f(s|N')} > 0. \tag{1}$$

The MLRP implies that a higher norm leads to stochastically higher signals (in the sense of first-order stochastic dominance), while a higher signal leads to a higher conditional distribution of the norm. This is shown in the following auxiliary lemma.

Lemma 1 *For any $s \in S$ and $N \in [0,1]$ we have*

$$\frac{\partial F(s|N)}{\partial N} < 0, \tag{2}$$

$$\frac{\partial G(N|s)}{\partial s} < 0. \tag{3}$$

Proof: See Appendix A. ■

Denote by E_{Ni} the expectation of agent i about the norm after observing s_i, i.e.,

$$E_{Ni} = E\left[N|s_i\right].$$

Lemma 1 straightforwardly implies that a higher private signal about the norm leads to a higher own expectation of the norm.

Corollary 1 *The expectation about the norm E_{Ni} strictly increases with s_i.*

[4] The same notation is used for conditional distribution/density functions.

Proof:

Using integration by parts, we obtain:

$$E_{Ni} = E[N|s_i] = \int_0^1 Ng(N|s_i)dN = 1 - \int_0^1 G(N|s_i)dN.$$

Then, the claim follows by Lemma 1. ■

Next, let us consider how the revealed information about the expectation of another agent affects own beliefs about the norm, and vice versa.

2.3 The effect of information about the expectation of another agent

Assume that agent i now additionally receives information on $j's$ beliefs about her own choice a_i. To be specific, let us assume that she directly observes $j's$ expectation $E_j = E_j\left[a_i|s_j\right]$ while the norm N remains unknown to both agents. As j does not have further information on i besides s_j this expectation must be equal to

$$E_j = E_j\left[a_i|s_j\right] = E_j\left[a|s_j\right] = E\left[N|s_j\right] = E_{Nj}. \tag{4}$$

Note that since E_{Nj} is a continuous and strictly increasing function of s_j by Corollary 1, agent i can perfectly infer s_j from observing E_{Nj} using the inverse function which we here denote as $s_j(E_{Nj})$. Then, agent's i posterior belief about the norm after observing both s_i and E_j is

$$E_i[N|s_i, E_j] = E_i[N|s_i, s_j] = \int_0^1 Ng(N|s_i, s_j)dN = 1 - \int_0^1 G(N|s_i, s_j)dN, \tag{5}$$

where to derive the last term we used integration by parts. Since s_i and s_j are independently distributed conditional on the norm N, the result of 1 still holds with respect to $G(N|s_i, s_j)$, i.e., $\frac{\partial G(N|s_i,s_j)}{\partial s_j} < 0$.[5] This together with (5) yields

$$\frac{\partial E_i[N|s_i, E_j]}{\partial s_j} > 0. \tag{6}$$

Finally, by the chain rule we obtain

$$\frac{\partial E_i[N|s_i, E_j]}{\partial E_j} = \frac{\partial E_i[N|s_i, E_{Nj}]}{\partial E_{Nj}} = \frac{\partial E_i[N|s_i, E_{Nj}]}{\partial s_j}\frac{\partial s_j(E_{Nj})}{\partial E_{Nj}} > 0, \tag{7}$$

[5] In particular, the MLRP also holds for the conditional distribution $F(s_j|N, s_i)$ since

$$\frac{\partial}{\partial s_j}\frac{f(s_j|N'')}{f(s_j|N')} = \frac{\partial}{\partial s_j}\frac{f(s_j|N'', s_i)}{f(s_j|N', s_i)}.$$

Hence, the arguments in the proof of 1 continue to hold.

where the first equality is by (4), while the inequality follows from (6) and Corollary 1. If the agent is not guilt averse ($\theta_{Gi} = 0$) but $\theta_{Ni} > \bar{\theta}_N$, she will thus choose $a_i = E_i[N|s_i, E_j]$ which is strictly increasing in E_j by (7).

Hence, we have shown the following result:

Proposition 1 *Under uncertainty about the norm, observing a higher expectation of another agent leads to a higher posterior belief about the norm. In turn, even if the agent is not guilt averse ($\theta_{Gi} = 0$), higher revealed expectations of the other agent lead to a higher action, i.e.*

$$\frac{\partial a_i}{\partial E_j} > 0$$

if the agent is sufficiently norm compliant ($\theta_{Ni} > \bar{\theta}_N$).

Proposition 1 thus implies that if an agent is a norm complier but not guilt averse then under uncertainty about the social norm she should react to disclosed expectations of others even though she does not care about these expectations *per se*.

2.4 The effect of information about the social norm

Let us show that private signal s_i about the social norm also affects i's second-order beliefs about E_j in case if the latter is unknown to i. The second-order belief of i about E_j is

$$\begin{aligned} E_i[E_j] &= E_i[E_j|s_i] = E_i[E_{Nj}|s_i] = E_i\left[\int_0^1 Ng(N|s_j)dN \,\middle|\, s_i\right] \\ &= \int_S \left(\int_0^1 Ng(N|s_j)dN\right) f(s_j|s_i)ds_j \qquad (8) \\ &= \int_S \left(1 - \int_0^1 G(N|s_j)dN\right) f(s_j|s_i)ds_j \\ &= 1 - \int_S \left(\int_0^1 G(N|s_j)dN\right) f(s_j|s_i)ds_j \\ &= 1 - \int_0^1 G(N|\bar{s})dN + \int_S \int_0^1 \frac{\partial G(N|s_j)}{\partial s_j} dN F(s_j|s_i)ds_j, \qquad (9) \end{aligned}$$

where we used integration by parts to obtain the third and the fifth equalities, and $\bar{s}$ is the upper bound of S. Consequently,

$$\frac{\partial E_i[E_j]}{\partial s_i} = \int_S \int_0^1 \frac{\partial G(N|s_j)}{\partial s_j} dN \frac{\partial F(s_j|s_i)}{\partial s_i} ds_j. \qquad (10)$$

At the same time, using the law of total probability and integration by parts, we obtain

$$\begin{aligned} F(s_j|s_i) &= \int_0^1 F(s_j|N, s_i)g(N|s_i)dN \\ &= \int_0^1 F(s_j|N)g(N|s_i)dN \\ &= F(s_j|1) - \int_0^1 \frac{\partial F(s_j|N)}{\partial N} G(N|s_i)dN. \end{aligned}$$

This together with Lemma 1 implies

$$\frac{\partial F(s_j|s_i)}{\partial s_i} < 0.$$

Substituting this into (10) and taking into account $\frac{\partial G(N|s_j)}{\partial s_j} < 0$ by Lemma 1, we finally have

$$\frac{\partial E_i[E_j]}{\partial s_i} > 0. \tag{11}$$

If the agent is not norm compliant ($\theta_{Ni} = 0$) but $\theta_{Gi} > \bar{\theta}_G$, she will thus choose $a_i = E_i[E_j]$ which is strictly increasing in s_i by (11). Hence, we obtain the following result.

Proposition 2 *Under uncertainty about the expectation of another agent, a higher signal about the norm leads to a higher second-order belief about this expectation. In turn, even if the agent is not norm compliant ($\theta_{Ni} = 0$), a higher signal about the norm leads to a higher action, i.e.*

$$\frac{\partial a_i}{\partial s_i} > 0$$

if the agent is sufficiently guilt averse ($\theta_{Gi} > \bar{\theta}_G$).

Analogously to Proposition 1, Proposition 2 implies that if an agent is purely guilt averse, then revealed information about the norm may change her behavior even though she might not directly care about compliance to social norms.

2.5 Experimental implications

Thus, we have formally shown a mutual signaling effect: revealed expectations of others signal information about the social norm (under uncertainty about the latter), and vice versa. Hence, in order to distinguish guilt aversion and norm conformity from each other, one needs to control for such indirect signaling effects. The most straightforward way to do it is to ensure that subjects are certain (in a reasonable approximation) about both norms and expectations. In this case, the mutual signaling effects, which rely on incomplete information, are trivially excluded (e.g., a revealed expectation of another subject, being itself a noisy signal about the norm, cannot significantly influence one's own belief about the norm anymore if one has already obtained a sufficiently precise signal about the norm). This is the approach we undertake in our experiment based on the dictator game, which is presented in the next sections.

3 Experimental Design

The experiment consisted of three parts. At the beginning of the experiment all participants received the instructions for Parts 1 and 2 and learned about their roles. The roles (either of dictator or recipient) were assigned randomly and remained unchanged until the end of the experiment. In Part 1 subjects had to provide their expectations regarding the dictator game following in Part 2. The underlying dictator game was as follows: The dictator had to allocate €14 between himself and one randomly assigned recipient. After reading the instructions for Parts 1 and 2, recipients had to estimate the average dictators' transfer in Part 2 (in their experimental session) and dictators were asked to state their beliefs about their matched recipient's guess of the average transfer, i.e. to submit their second-order beliefs. Subjects received €5 for their answer that deviated from the true value by no more than €0.15. The initial endowment of each subject (i.e., the amount obtained independently of dictators' allocation decisions) was €7.50 in Part 1, and €2.50 in Part 2.

After the first two parts were finished, subjects were informed about whether their guess from Part 1 earned the bonus.[6] After that instructions for Part 3 were distributed, while each dictator was matched with a new recipient. Each recipient was then asked whether she agrees that her expectation, elicited in Part 1, can be transmitted to dictators. If the recipient agreed, she obtained an additional endowment of €2.50. This procedure was used to avoid experimental deception by omission: Recipients became fully aware and in control of their belief disclosure to dictators before Part 3 started, while - at the same time - they couldn't strategically distort their guesses in Part 1 (see Khalmetski et al., 2015, for a discussion). Dictators always received an initial endowment of €2.50 for Part 3.[7]

Part 3 is now our main stage of interest. Also here, dictators had to decide how to allocate another €14 between themselves and their newly matched recipient. Before this decision, they could see two pieces of information: (i) the average transfer that *the other* dictators in the same session made in Part 1[8] and (ii) the guess of the matched recipient about the average transfer of the dictators.[9] The information was presented to dictators in a random order. After dictators made their decision in Part 3, one part

[6] This was done to gain experimental control over subjects' beliefs regarding the total payoff from the experiment. Controlling for this information in our regression analysis doesn't change results in any way.

[7] When recipients agreed to transmit their expectations, the endowments became equivalent to the ones in Part 2.

[8] If, for instance, the session size was 16 subjects, the descriptive norm was the average transfer of the 7 other dictators. Hence, we use the natural sampling variation of the average transfers among the sessions to identify the causal effect of descriptive norms on behavior.

[9] This information was withheld if recipients did not grant the permission.

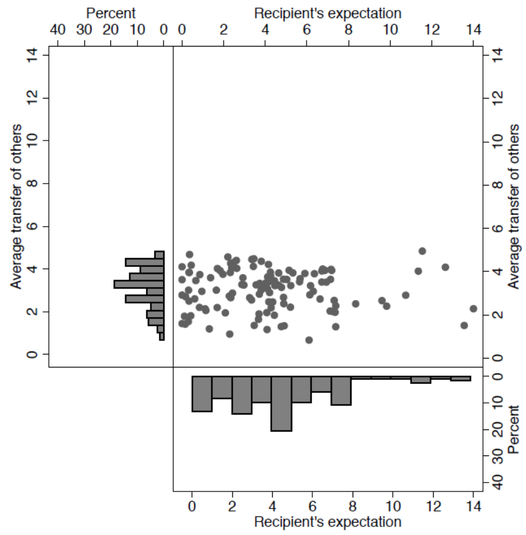


Figure 1: Distribution of descriptive norms and recipient's expectations.

was randomly chosen for payment and subjects obtained their earnings in cash.

The experiment was conducted in the Cologne Laboratory for Economic Research with 248 participants. Participants were recruited with ORSEE (Greiner, 2015), and the experiments were computerized with z-Tree (Fischbacher, 2007). In each experiment we ran 16 sessions of either 14 or 16 subjects. The variation was caused by different attendance rates. The instructions were distributed on paper and can be found in Appendix C. The average earning was around €9 (including the show-up fee), and the experiment lasted for about 45 minutes.

4 Results

As a first step it is important to see whether the heterogeneity in individual transfers generated sufficient variation in the descriptive norms displayed to dictators. Across all sessions, the average dictator transfer in Part 2 was €3.12 (SD = 2.70). The *average transfer of the others* that was displayed to dictators in Part 3 varied between 0.86 and 4.82 (Mean = 3.12, SD = 0.92).[10] In Part 3, 121 out of 124 recipients agreed to transmit their guesses to dictators so that we obtain 121 observations for the analysis, while avoiding substantial selection effects. The transmitted recipient's expectation varied between €0 and €13.85 (Mean = 4.17, SD = 2.98). Around 90% of the expectations were below or equal to the half of the pie. Figure 1 gives an overview of the distribution of both benchmarks.

The average transfer in Part 3 was 2.56 (SD = 2.32). This is significantly lower than in Part 2 ($p < 0.01$, Wilcoxon signed-rank test). On average, dictators reduced their

[10]Note that this is well in line with the natural variation arising from an ex-post Monte Carlo simulation on the basis of our data, see Appendix B.

Table 1: The effect of descriptive norm and recipient's expectation on transfers.

	(1) All	(2) Expect. ≤ 7	(3) $0.86 \leq$ Expect. ≤ 4.82
Unconditional transfer (in Part 2)	0.651^{***}	0.639^{***}	0.579^{***}
	(0.0409)	(0.0449)	(0.0599)
Norm (average transfer)	0.371^{***}	0.376^{**}	0.524^{***}
	(0.141)	(0.154)	(0.185)
Recipient's expectation	0.0310	0.0781	0.344^{**}
	(0.0445)	(0.0706)	(0.147)
Age	0.0848	0.0734	0.113^{*}
	(0.0522)	(0.0541)	(0.0662)
Gender	0.259	0.216	0.195
	(0.260)	(0.279)	(0.363)
Observations	121	109	64
Pseudo R^2	0.247	0.229	0.224

Tobit regressions; marginal effects reported; robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

transfers by €0.57 (SD = 1.67). Approximately half of dictators in our sample did not change their allocation decisions in Part 3 relative to Part 2 (50.8%). Those dictators who decreased their transfers (36.3%) transferred on average 41.73% less. Around 13% of subjects transferred more in Part 3 as compared to Part 2 (their transfers were 2.6 higher in Part 3 than in Part 2). Finally, the average dictator's second-order belief was €4.32 (SD = 3.48), with 89.5% of the beliefs being below or equal to the half of the pie.

Exploiting the exogenous variation in descriptive norms and recipient expectations, we can now estimate their causal effect on dictators' transfers. Table 1 shows the results (marginal effects) of Tobit regressions with the dictator transfer in Part 3 as the dependent variable and the descriptive norm and the expectation of the matched recipient as independent variables. Besides, we control for the dictator's own transfer in Part 2 as a measure for general social preferences, and for standard demographic variables such as age and gender.

We find strong evidence in favor of norm conformity: As column (1) shows, the information about the social norm has a sizable and significant effect on dictators' transfers. A one unit increase in the displayed norm increases dictators' transfers by 0.371 units.

At the same time, column (1) seems to indicate that the recipient's expectation does not affect transfers. In turn, one may be tempted to conclude that guilt aversion plays no role when individuals have information about the norm, which would be in line with the view that recipients' expectations affect behavior only as they signal information about the social norm. However, a closer inspection of the data suggests that this is not the appropriate explanation.

First of all, it is important to note that previous related experiments (Ellingsen et al., 2010; Khalmetski et al., 2015) in which subjects did not have information about the norm also did not detect aggregate effects of the recipient's (exogenously disclosed) expectations. One potential reason – as argued recently by Balafoutas and Fornwagner (2017) – is that guilt aversion may matter only if the expectations of the recipient are deemed "acceptable" or "legitimate". Balafoutas and Fornwagner (2017), for instance, show that expectations raise transfers in the dictator game only if they do not exceed a certain (individual specific) level which is typically close to the equal split (i.e. 7 in our context). Moreover, "unreasonably" high expectations beyond that level may even be "punished" with lower transfers. In our data, the recipient's expectations vary across a much larger interval (i.e. between 0 and 13.85) than the norm (between 0.86 and 4.82). Hence, it is possible that observed expectations are more likely to appear in an "unreasonable" range than the norm. To account for this, we first reduce the sample to the observations where the expectation does not exceed the equal split of 7.[11] As shown in column (2) the point estimate of the effect of recipient's expectation increases, but still remains insignificant. But when we further restrict the range and consider only expectations that lie in the interval spanned by the norm (i.e. between 0.86 and 4.82), the coefficient of the recipient's expectations becomes sizable and significant (column (3)). Hence, in the presence of information about the average behavior of others, the dictators react only to modest expectations of recipients.

This leads to two questions: First, can we say more about which expectations dictators deem acceptable? And second, how do dictators react to expectations beyond the acceptable range? In particular, there may be different individual benchmarks in our context to which an expectation could be compared by dictators. Among these benchmarks may be the equal split, the dictator's own unconditional transfer in the previous Part 2, the dictator's second-order belief about the expectation of the recipient (elicited in Part 1), and the displayed social norm. Table 2 shows regressions of the form

$$\begin{aligned} transfer_i &= \alpha + \beta \cdot transfer_i^0 + \gamma \cdot norm_i + \delta \cdot expectation_i \\ &\quad + \phi \cdot (expectation_i - benchmark_i) \cdot I_{\{expectation_i \geq benchmark_i\}} \end{aligned}$$

where $transfer_i$ denotes transfer in Part 3, $transfer_i^0$ is the (unconditional) transfer under no information in Part 2, and $I_{\{expectation_i \geq benchmark_i\}}$ is a dummy variable indicating whether the expectation exceeds the respective benchmark. In this way, we estimate piece-wise linear and continuous reaction functions (with respect to a change in the recipient's expectation). Here, ϕ estimates a potential change in the slope of

[11] Note that the expectation of the recipient is a purely exogenous variable from the perspective of the dictator, and hence such sample restriction does not reduce the statistical validity of the regression analysis.

Table 2: The effect of different reference standards on guilt aversion.

	(1)	(2)	(3)	(4)	(5)
Unconditional transfer (in Part 2)	0.652***	0.429***	0.644***	0.650***	0.401***
	(0.0401)	(0.0833)	(0.0450)	(0.0402)	(0.0887)
Norm (average transfer)	0.374***	0.355***	0.366**	0.380*	0.464**
	(0.142)	(0.135)	(0.143)	(0.197)	(0.216)
Recipient's expectation	0.0739	0.264***	0.0576	0.0206	0.138
	(0.0668)	(0.0864)	(0.0740)	(0.153)	(0.164)
(Recip. expectation - 7)	-0.157				-0.0485
$\times\ I_{\{Recip.\ expectation\ \geq\ 7\}}$	(0.113)				(0.154)
(Recip. expectation - Uncond. transfer)		-0.389***			-0.455***
$\times\ I_{\{Recip.\ expectation\ \geq\ Uncond.transfer\}}$		(0.101)			(0.126)
(Recip. expectation - SOB)			-0.0391		0.0743
$\times\ I_{\{Recip.\ expectation\ \geq\ SOB\}}$			(0.0708)		(0.0767)
(Recip. expectation - Norm)				0.0144	0.179
$\times\ I_{\{Recip.\ expectation\ \geq\ Norm\}}$				(0.179)	(0.224)
Age	0.0795	0.0582	0.0823	0.0852	0.0607
	(0.0514)	(0.0457)	(0.0520)	(0.0516)	(0.0446)
Gender	0.241	0.152	0.228	0.263	0.248
	(0.256)	(0.245)	(0.261)	(0.277)	(0.266)
Observations	121	121	121	121	121
Pseudo R^2	0.250	0.271	0.248	0.247	0.275

Tobit regressions; marginal effects reported; robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

the reaction function at the respective benchmark.[12] We compare different potential benchmarks (i.e. equal split of 7, previous transfer in Part 2, dictator's own second-order belief, and the displayed descriptive norm). Each model thus estimates whether the reaction function displays a kink at the relevant benchmark, i.e. whether its slope is significantly different above as compared to below the benchmark.

Table 2 shows the results of the regression analysis. As column (2) indicates, there is a kink in the dictators' reaction function at her own prior transfer choice in Part 2 (i.e. the transfer chosen before receiving information about the norm and the recipient's expectation). Below this threshold transfers are increasing in the recipient's expectations, but above this threshold they are strictly decreasing. As a Wald test confirms,

[12] Indeed, assume that the true data generating process for some dependent variable y and regressor x is of the form:

$$y_i = \begin{cases} \alpha + \beta_1 x_i + \varepsilon_i \text{ if } x_i < b_i, \\ \alpha + \beta_1 b_i + \beta_2 (x_i - b_i) + \varepsilon_i \text{ if } x_i \geq b_i, \end{cases}$$

with $\beta_2 \neq \beta_1$ and b being some individual benchmark. Thus, y as a function of x has a kink at the benchmark. One can easily verify that this expression for y_i is equivalent to

$$y_i = \alpha + \beta_1 x_i + (\beta_2 - \beta_1)(x_i - b_i) I_{\{x_i \geq b_i\}} + \varepsilon_i$$

so that the resulting coefficient on $x_i I_{\{x_i \geq b_i\}}$ measures the change in the slope at the benchmark.

the estimated slope beyond the kink of $0.264 - 0.389 = -0.125$ is significantly smaller than zero ($p = 0.012$). We find no evidence that there is a structural break at any of the other potential benchmarks. Importantly, the effect of displayed descriptive norm remains statistically significant in all specifications.

As another way to study how the dictators react to the information about the norm and the recipient's expectation, we measure how dictators update their transfers in Part 3 (relatively to the unconditional transfer in Part 2) after learning about the norm and the recipient's expectation. For this, we regress the difference in transfers between Part 3 and Part 2 on the dummy variables measuring how much the observed information (either about the norm or the recipient's expectation) deviates from their unconditional transfers in Part 2. Thus, we run the following specifications allowing for a non-linear reaction to the novel information contained in the norm and the recipient's expectation:

$$\begin{aligned}\Delta transfer_i &= \alpha + \gamma_1 I_{\{\Delta_i^{norm} < -3\}} + \gamma_2 I_{\{\Delta_i^{norm} \in [-3,-1]\}} + \gamma_3 I_{\{\Delta_i^{norm} \in [1,3]\}} \\ &\quad + \gamma_4 I_{\{\Delta_i^{norm} > 3\}} + \delta_1 I_{\{\Delta_i^{exp.} < -3\}} + \delta_2 I_{\{\Delta_i^{exp.} \in [-3,-1]\}} \\ &\quad + \delta_3 I_{\{\Delta_i^{exp.} \in [1,3]\}} + \delta_4 I_{\{\Delta_i^{exp.} > 3\}} + \varepsilon_i\end{aligned}$$

where $\Delta_i^{norm} = norm_i - transfer_i^0$ and $\Delta_i^{exp.} = expectation_i - transfer_i^0$. The interval $(-1, 1)$ is taken as baseline. Hence, the coefficients γ_k (δ_k), $k = 1, ..., 4$, reflect how much the dictator updates her transfer in Part 3 after learning that her prior transfer deviated from the norm (expectation) by the corresponding amount, relative to the case where the observed norm (expectation) was close to the prior transfer. Figure 2 plots the coefficients for the considered intervals and their 95% confidence bands.

The right panel shows the dictator's reaction to information about the recipient's expectation that differs from her own previous transfer. It confirms the pattern detected in our piece-wise linear regression in the above: Dictators reduce their transfers when learning that their recipient expects less that the previously given, but do not increase their transfers if the recipient expects more (even "punishing" too high expectations with lower transfers). In contrast, dictators react monotonically to information about the norm as the left panel shows: they reduce their transfers when the norm is lower than their own prior transfer, and tend to increase their transfers if it is above.

Hence, when both the descriptive social norm and the expectation of the recipient are known, (i) dictator transfers are increasing in the social norm and (ii) dictators also react to the recipient's expectation but only if this expectation falls below their own prior choice. In other words, information about the social norm pushes transfers both ways, but information about the recipient's expectation is used in a self-serving manner to reduce one's own giving.

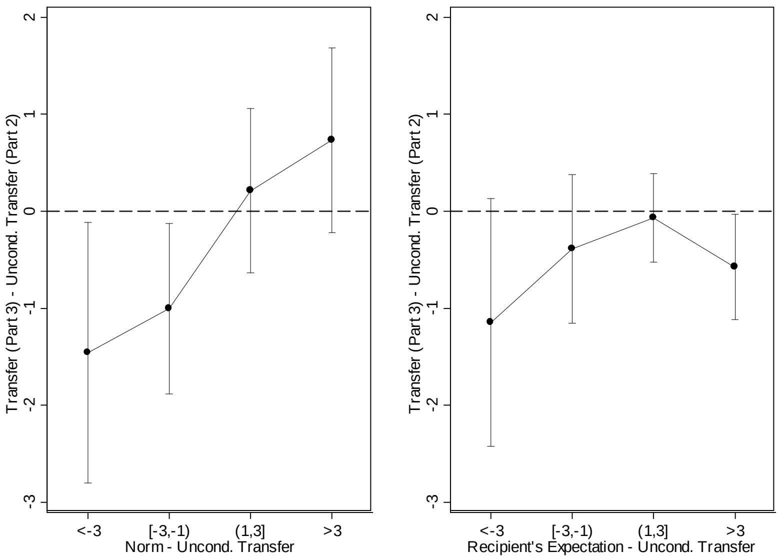


Figure 2: Reaction to information about the norm (left panel) and the recipient's expectation (right panel).

5 Conclusion

We find that both (descriptive) social norms and revealed expectations of others affect behavior even if one controls for the fact that both of these types of information may mutually signal each other. Notably, dictators tend to ignore whether the recipient's expectation is above or below the displayed norm, but rather use their own unconditional transfer as a benchmark to assess the reasonableness of the expectation applying it in a self-serving manner.

Our results on non-monotonicity of guilt aversion comply with the previous related evidence of Pelligra et al. (2016), Khalmetski (2016) and Balafoutas and Fornwagner (2017), and motivate a refinement of the notion of guilt aversion formalized by Battigalli and Dufwenberg (2007). In particular, it might be reasonable to assume that not all expectations of other players equally matter for the utility of a (guilt averse) decision maker while those which exceed one's own counterfactual behavior (under prior beliefs) are downweighted. Our results could also refine these of Khalmetski et al. (2015) who found that in order to measure the actual effect of guilt aversion one needs to take into account the heterogeneity in belief-dependent preferences (such as eagerness to positively surprise others). The current study further complements it by showing that the heterogeneity in the revealed expectations should also be taken into account.

Overall, our results point out that both norm conformity and guilt aversion (for the reasonable range of expectations) are important in shaping individual decisions.

References

Balafoutas, L. and H. Fornwagner (2017). The limits of guilt. *Journal of the Economic Science Association 3*(2), 137–148.

Bardsley, N. and R. Sausgruber (2005). Conformity and reciprocity in public good provision. *Journal of Economic Psychology 26*(5), 664 – 681.

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

Bénabou, R. and J. Tirole (2012). Laws and norms. *IZA Discussion Paper 6290*.

Bicchieri, C. and E. Xiao (2009). Do the right thing: but only if others do so. *Journal of Behavioral Decision Making 22*(2), 191–208.

Charness, G. and M. Dufwenberg (2006). Promises and partnership. *Econometrica 74*(6), 1579–1601.

Cialdini, R. B., R. R. Reno, and C. A. Kallgren (1990). A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology 58*(6), 1015.

Danilov, A. and D. Sliwka (2016). Can contracts signal social norms? experimental evidence. *Management Science 63*(2), 459–476.

Dufwenberg, M. and U. Gneezy (2000). Measuring beliefs in an experimental lost wallet game. *Games and economic Behavior 30*(2), 163–182.

Ellingsen, T., M. Johannesson, S. Tjøtta, and G. Torsvik (2010). Testing guilt aversion. *Games and Economic Behavior 68*(1), 95–107.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics 10*(2), 171–178.

Friebel, G. and W. Schnedler (2011). Team governance: Empowerment or hierarchical control. *Journal of Economic Behavior & Organization 78*(1), 1–13.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association 1*(1), 114–125.

Hauge, K. E. (2016). Generosity and guilt: The role of beliefs and moral standards of others. *Journal of Economic Psychology 54*, 35–43.

Khalmetski, K. (2016). Testing guilt aversion with an exogenous shift in beliefs. *Games and Economic Behavior 97*, 110–119.

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

Krupka, E. L., S. Leider, and M. Jiang (2017). A meeting of the minds: Informal agreements and social norms. *Management Science 63*(6), 1708–1729.

Krupka, E. L. and R. A. Weber (2009). The focusing and informational effects of norms on pro-social behavior. *Journal of Economic psychology 30*(3), 307–320.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics 12*(2), 380–391.

Pelligra, V., T. Reggiani, and D. J. Zizzo (2016). Responding to (un) reasonable requests. Technical report.

Sliwka, D. (2007). Trust as a signal of a social norm and the hidden costs of incentive schemes. *The American Economic Review 97*(3), 999–1012.

Van der Weele, J. (2009). The signaling power of sanctions in social dilemmas. *The Journal of Law, Economics, & Organization 28*(1), 103–126.

Appendix A: Proof of Lemma 1

The first inequality follows from Proposition 2 in Milgrom (1981). Let us show the second inequality. Fix any $N' \in [0,1]$. We have

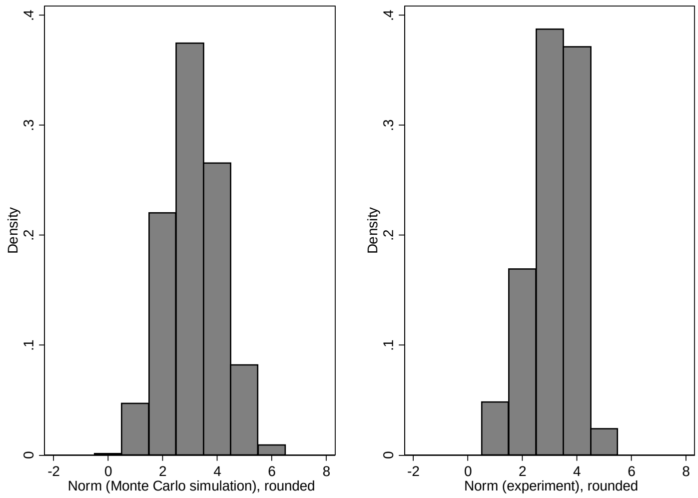
$$\begin{aligned} G(N'|s) &= \Pr[N \leq N'|s] \\ &= \frac{f[s|N \leq N']G(N')}{f[s|N \leq N']G(N') + f[s|N > N'](1-G(N'))} \\ &= \frac{1}{1+\frac{f[s|N>N'](1-G(N'))}{f[s|N\leq N']G(N')}} = \frac{1}{1+\frac{(1-G(N'))}{G(N')}\gamma(s)}, \end{aligned} \tag{12}$$

where the second equality is by Bayes rule, and $\gamma(s) \equiv \frac{f[s|N>N']}{f[s|N\leq N']}$. Then, by the law of total probability

$$\begin{aligned} \gamma(s) &= \frac{f[s|N > N']}{f[s|N \leq N']} = \frac{\int_{N'}^{1} f(s|x)g(x|x > N')dx}{\int_{0}^{N'} f(s|y)g(y|y \leq N')dy} \\ &= \frac{\frac{1}{1-G(N')}\int_{N'}^{1} f(s|x)g(x)dx}{\frac{1}{G(N')}\int_{0}^{N'} f(s|y)g(y)dy} \\ &= \frac{G(N')}{1-G(N')}\int_{N'}^{1} \frac{f(s|x)}{\int_{0}^{N'} f(s|y)g(y)dy} g(x)dx \\ &= \frac{G(N')}{1-G(N')}\int_{N'}^{1} \frac{1}{\int_{0}^{N'} \frac{f(s|y)}{f(s|x)} g(y)dy} g(x)dx. \end{aligned}$$

Since $x > y$ for any x and y under the integral signs (except for $x = y = N'$), (1) implies that $\frac{f(s|y)}{f(s|x)}$ strictly decreases with s. Consequently, the whole function $\gamma(s)$ strictly increases with s. Then, by (12) we obtain that $G(N'|s)$ strictly decreases with s. ∎

6 Appendix B: Sampling distribution of average transfers



Appendix C: Experimental instructions

General information

Welcome to the experiment! The goal of this experiment is to study individual behavior in particular situations. If you have a question, please raise your hand. We will be glad to help you at your seat. **During the experiment, any other communication is not permitted!**

In this experiment, you can earn money. How much you earn depends on your decisions as well as on the decisions of the other participants. More detailed information about this is provided in the experimental instructions.

Your payoff will be paid to you personally in cash at the end of the experiment.

Your payoff and your decisions will be treated strictly confidentially. None of the participants will get to know during or after the experiment with whom he interacted. Your decisions are hence **anonymous**.

Experiment

This experiment consists of **three** parts.

Your payoff and the payoffs of the other participants are obtained from one of the three parts. This means that at the end of the experiment **one part will be randomly selected for all participants and paid out.**

Thus, thoroughly consider your decisions in each part of the experiment! Any of your decisions may result in a monetary payoff and therefore influence your today's income.

Next, you will receive instructions for the first and the second parts of the experiment. After the second part is over, you will receive instructions for the third part of the experiment.

Part 1 and Part 2

All participants are randomly divided into participants A and participants B. Every participant is matched to another person in the other role, so that **each participant A is matched to one participant B**. Both participants are seated in this room. The assignment of roles and the matching of participants to each other stays the same in part 1 and part 2. You will see on the computer screen which role you are assigned to.

Part 1

As described above, the earnings from part 1 will be paid out to all participants at the end of the experiment with probability 1/3 (otherwise, the earnings from part 2 or part 3 will be paid out).

In part 1, every participant receives a show-up fee of €2.50.

In addition to this, every participant receives an endowment of €5.

The task of all participants in part 1 is to guess the behavior of other participants *in part 2* of the experiment as precisely as possible. Every participant can earn an additional payoff by a good guess. You will find further information on this on your screen, after the rules for part 2 are explained.

The participants will be informed whether their guess has earned an additional payoff after the second part of the experiment.

Part 2

As described above, the earnings from part 2 will be paid out to all participants at the end of the experiment with probability 1/3 (otherwise, the earnings from part 1 or part 3 will be paid out).

In part 2, every participant receives a show-up fee of €2.50.

Decision of participant A:

Participant A receives an endowment of €14. He can give a part of his endowment to participant B.

Decision of participant B:

Participant B does not take any decision about the division of the endowment.

Therefore, the **payoffs** are calculated as follows:

Payoff to participant A = €14 − amount given to participant B

Payoff to participant B = Amount given by participant A

Participant B will be informed about the amount that was given to him by participant A only at the end of the experiment, namely after the third experimental part.

This is the end of the instructions for parts 1 and 2. Please take your time, and be sure to understand these instructions. **If you have any questions, raise your hand and an experimenter will come to you.**

Part 3

In the third part of the experiment, all participants retain their roles (participant A or participant B) which were previously assigned to them.

For every participant A, a random mechanism will select one participant B from this room who has not interacted with the participant A in the previous parts of the experiment. This person will be the recipient of the amount that participant A gives in part 3.

As described above, the earnings from part 3 will be paid out to all participants at the end of the experiment with probability 1/3 (otherwise, the earnings from part 1 or part 2 will be paid out).

In part 3, participant A receives an amount of €2.50.

At the beginning of part 3, participants B can decide whether or not their guess submitted in the first part (regarding the average amount which participants A send to participants B) may be transmitted to participant A. For the disclosure of this information participants B receive an amount of €2.50.

Decision of participant A:

As in part 2, participant A receives an endowment of €14. He can give a part of his endowment to the participant B who is now matched to him.

Decision of participant B:

Participant B does not take any decision about the division of the endowment.

Additionally, participant A receives the following information:

a) (If participant B has agreed to transmit his/her guess) The expectation of the currently matched participant B (namely, the recipient of the amount given in part 3) about the average amount which was given by participants A to participants B in part 2.

b) The average amount which was given by participants A from this room to participants B in part 2.

Therefore, the **payoffs** are calculated as follows:

Payoff to participant A = €14 − amount given to participant B

Payoff to participant B = Amount given by participant A

Participant B will be informed about the amount that was given to him by participant A at the end of the experiment.

This is the end of the instructions for the third part. Please take your time, and be sure to understand these instructions. **If you have any questions, raise your hand and an experimenter will come to you.**

[*Belief elicitation questions (shown on screen):*]

[*For participants B:*]

In what follows, you will be asked a question to provide a guess. If your guess to this question does not deviate from the true value by more than 15 Cent, you will get a bonus of €5. The question refers to the behavior of the participants in this room. Please try to answer the question as good as possible.

Question:

Please guess the **average** amount which will be given by participants A to participants B in part 2 of the experiment.

The average amount which participants A give to participants B is (from 0.00 to 14.00 Euro):

[*For participants A:*]

In what follows, you will be asked a question to provide a guess. If your guess to this question does not deviate from the true value by more than 15 Cent, you will get a bonus of €5. The question refers to the behavior of the participants in this room. Please try to answer the question as good as possible.

We have asked participants B the following question:

*"Please guess the **average** amount which will be given by participants A to participants B in part 2 of the experiment."*

Also participants B get a bonus of €5 for a guess which does not deviate from the true value by more than 15 Cent.

Question:

Please guess the answer to this question of the participant B **who is matched to you** (namely, what do you think is the belief of the participant B who is matched to you about the average amount given by participants A?).

The amount which is expected by participant B is (from 0.00 to 14.00 Euro):