INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Exogenous Variation of Endowments and Beliefs in a Dictator Game

Pierpaolo Battigalli, Giovanni Di Bartolomeo
and Stefano Papa

Working Paper n. 725

This Version: March 2026

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Exogenous Variation of Endowments and Beliefs in a Dictator Game*

Pierpaolo Battigalli† Giovanni Di Bartolomeo‡ Stefano Papa§

March 2026

Abstract

We study theoretically and experimentally sharing behavior in a dictator game with asymmetrically observed random initial allocation of a fixed surplus. Our findings are consistent with two other-regarding channels mitigating selfish behavior: a belief-independent desire to comply with an equal-sharing norm, and belief-dependent guilt aversion. The norm-channel allows us to manipulate first and second-order beliefs by changing the commonly known distribution of the initial allocation observed only by the dictator. Comparing behavior for fixed endowment in the high and low-expectations treatments, we test for the presence of guilt aversion.

Keywords: sharing norm, opportunity cost, first and second-order beliefs, guilt aversion.

JEL classification codes: A13; C91; D01; D64.

*We thank Carlo Andreatta, Sem Manna, Elena Manzoni, and Silvia Meneghesso for useful comments. Our experiment was designed in 2021 and conducted in 2022 at Tor Vergata University (Rome.) The hypotheses of the experiment were not pre-registered. However, our hypotheses were articulated before the experiment. An earlier version of this paper was circulated under the title "Guilt, Inequality, and Gender in a Dictator Game."

†Department of Decision Sciences, IGIER, and BIDSA, Bocconi University, Italy; pierpaolo.battigalli@unibocconi.it.

‡Department of Economics and Law and CIMEO, Sapienza University of Rome, Italy; Department of Economics, University of Antwerp, Belgium; giovanni.dibartolomeo@uniroma1.it.

§Department of Economics and Finance, University of Rome Tor Vergata, Italy; CIMEO, Sapienza University of Rome, Italy; stefano.papa@uniroma2.it.

1 Introduction

Dictator games have long been a workhorse for studying prosocial behavior. Two broad motivations dominate the literature. The first is belief-independent and encompasses norm-based and distributional preferences—such as altruism and fairness—according to which individuals share because they value others' payoffs or because they wish to comply with an equal-split norm, regardless of what the recipient expects (Fehr and Schmidt, 1999, 2010; Gächter *et al.*, 2017).[1] The second is belief-dependent: under guilt aversion, dictators experience a psychological cost when they fail to meet what they believe the recipient expects, so generosity depends on second-order beliefs about the recipient's expectations (Battigalli and Dufwenberg, 2007).[2]

Empirically testing these motivations is inherently challenging, because beliefs and behavior often co-move ex post, making it difficult to determine whether beliefs causally drive generosity or simply reflect it.[3] Yet the distinction is important: norm-based motives and belief-based motives have different implications across social contexts and, therefore, for policy design (see, e.g., Rothenhäusler *et al.*, 2018).

We address this challenge by combining a simple theoretical framework with a controlled experimental design that generates exogenous variation in the belief environment and exogenous variation in the realized opportunity cost of sharing. The theory highlights two channels. A social-norm channel links the opportunity cost of sharing to what is considered appropriate, shaping what participants expect dictators to do. A guilt-aversion channel predicts that dictators' behavior responds to second-order beliefs, especially when keeping is expected to impose a larger loss on the recipient.

Guided by our theoretical predictions, we implement a modified dictator game

[1] For a detailed taxonomy and definition of social norms, see Krupka and Weber (2013). The concept of social norms may also overlap with belief-dependent motivations (e.g., d'Adda *et al.*, 2020).

[2] Battigalli and Dufwenberg (2007) proposed a portable model of belief-based guilt aversion, later extended to allow for incomplete information and non-equilibrium analysis (e.g., Attanasi *et al.*, 2016, 2025; Battigalli *et al.*, 2019).

[3] Beliefs and actions are often ex post correlated, regardless of which motivation drives generosity (Khalmetski, 2016). To test these motivations, methods introducing exogenous variation in expectations are necessary (Vanberg, 2008; Khalmetski, 2016; Di Bartolomeo *et al.*, 2019).

with random and asymmetrically observed endowments. The dictator's endowment is realized after belief elicitation, and the recipient does not observe the realized endowment.[4] We manipulate the publicly known probability of facing a low- versus high-endowment game, which shifts first- and second-order beliefs about sharing, while the realized endowment determines the material cost of sharing. We run the experiment with 512 undergraduate participants at the University of Rome Tor Vergata.

Our work relates to evidence that endowments and opportunity costs matter for giving. Korenok *et al.* (2012, 2014) show that endowment size affects dictators' choices and that refraining from taking is not equivalent to giving. Viet Nguyen (2019) documents that dictators share less when others receive much larger endowments, pointing to self-serving fairness considerations. Relatedly, Di Bartolomeo and Papa (2016), building on Cox (2004), suggest that conditional motives can be more prominent when the opportunity cost of prosocial behavior is low.

We are also closely connected to the literature on guilt aversion in dictator games. Ellingsen *et al.* (2010) and Khalmetski *et al.* (2015) stress both the conceptual difficulty of distinguishing models and the practical challenges of belief elicitation. Methodologically, we build on Di Bartolomeo *et al.* (2019), who generate exogenous variation in expectations using asymmetric information and an experimental mechanism that shifts what players anticipate about others' behavior.[5]

The interaction between guilt aversion and social norms in dictator games is studied by Danilov *et al.* (2021) with a different experiment driven by a different model. While we assume that dictators may be driven by guilt aversion or an injunctive (equal splitting) norm, they assume that dictators may be guilt averse or want to comply with a descriptive norm, i.e., they may want to be close to average behavior. They experimentally disentangle descriptive-norm compliance and guilt aversion by giving dictators information about average behavior and about recipients' expectations.

The rest of the paper is organized as follows. Section 2 introduces the theo-

[4] Dictator-game giving is known to respond to exogenous changes in observability and the social or informational environment (Servátka, 2009).

[5] This methodology was used also in Ciccarone *et al.* (2020) and Di Bartolomeo *et al.* (2024). See Sections 2 and 3 for the details of our experimental design.

retical framework and the hypotheses guiding the experimental design. Section 3 describes the experimental implementation and the sources of variation used to separate belief-dependent motives from material incentives. Section 4 presents the analysis and main findings. Section 5 reports additional exploratory results on gender heterogeneity. Section 6 concludes. The Appendix contains details about belief elicitation and provides conditions on the primitives of our model that yield an auxiliary assumption of Section 2.

2 Experimental design and hypotheses

In this section, we explain our experimental design (2.1), we put forward a simple theoretical model with other-regarding motivations (2.2), and we present our theory-driven experimental hypotheses (2.3).

2.1 A dictator game with hidden initial allocations

We analyze a *Share-or-Keep Dictator Game* with random initial endowments (Battigalli *et al.*, 2025). The game is illustrated in Figure 1, where payoffs represent monetary payments rather than utilities, acknowledging that individual decisions might be influenced by other-regarding preferences, such as social norms and feelings of guilt. In this setup, given the observed realization $x \in \{x_L, x_H\}$ (with $x_L < x_H$) of a random initial allocation $\mathbf{x}$ for the dictator and $M - \mathbf{x}$ for the recipient, the dictator faces a binary decision: to *Keep* the initial allocation, preserving the status quo, or to *Share*, redistributing the resources equally between the two parties.

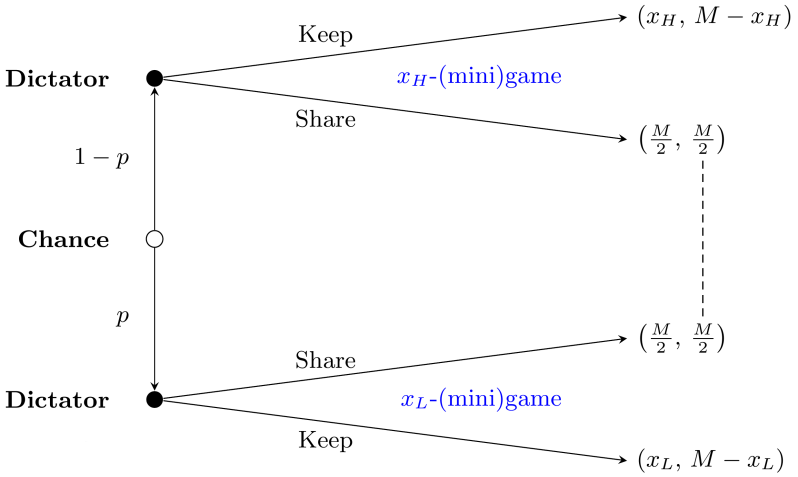

Figure 1 – Mini-Dictator Game with Random Initial Endowments.

Specifically, Chance splits M tokens as endowments of the dictator and the recipient, assigning x_L or x_H tokens to the dictator with, respectively, probability p or $1-p$, and the rest to the recipient. Next, the dictator can either *Share* the total number of tokens (M) equally or *Keep* the initial allocation (status quo). The distribution of the random endowment $\mathbf{x}$ is commonly known, but its realization x is observed *only* by the dictator (asymmetric information), and this is common knowledge. The dotted line emphasizes that if the dictator shares, the recipient cannot infer the dictator's initial endowment from observing their own payoffs.[6]

Of course, the own monetary payoff maximizing decision of the dictator is to always keep, regardless of the observed value x. However, incorporating other-regarding preferences shifts the focus to how social norms and anticipated guilt from keeping might shape the dictator's decision. The material opportunity cost of sharing, $x - M/2$, increases linearly in x. Therefore, with a slight abuse of language, we often refer to x as the (opportunity) cost of sharing. Furthermore,

[6] This reduces the potential impact of self-image concerns (Benabou and Tirole, 2006; Murnighan *et al.*, 2001).

for the sake of brevity, we refer to the "subgame" where the dictator initially holds x_H as the x_H-Game, while the other is called the x_L-Game.[7]

Our **experimental design** involves two treatments, which differ in the value of the probability $p = \Pr(\mathbf{x} = x_L)$ of playing the low-endowment (*low-cost*) game. In **treatment 1** (t_1), we set $p = 0.75$. In **treatment 2** (t_2), we set $p = 0.25$. The other parameters are: $M = 40$, $x_L = 24 > M/2$, and $x_H = 40 = M$.

To anticipate, our main experimental hypotheses are that (a) sharing is more frequent and recipients' first-order beliefs are higher in t_1 than in t_2, and (b) dictators' second-order beliefs reflect this.

2.2 Social norms and guilt aversion: A simple model

We assume that a fraction $q_n \in (0, 1)$ of agents in the dictator role, referred to as "norm-driven (n-)dictators," dislike violating the equal-sharing norm. The remaining fraction $q_g = 1 - q_n$ of agents are "guilt-driven (g-)dictators," who are instead motivated by guilt aversion. Furthermore, we assume that within each subgroup, agents are heterogeneous, i.e., they differ in the degree to which they care about adhering to social norms or about guilt.[8]

The parameterized utility function of an n-dictator is

$$u_n(m_D, m_R) = m_D - \theta_n \mathbf{1}_{m_D > m_R}(m_D, m_R), \tag{1}$$

where m_i is the realized monetary payoff of agent i in $\{D, R\}$,[9] $\mathbf{1}_{m_D > m_R}$ is the indicator function of (dictator-biased) unequal distribution, and parameter $\theta_n \geq 0$ gauges the psychological cost of violating the social norm of equal sharing.

The parameterized utility function of a g-dictator is

$$u_g(m_D, m_R, \alpha_R) = m_D - \theta_g \left[\mathrm{E}_\alpha(\mathbf{m}_R) - m_R\right]^+, \tag{2}$$

[7] Technically, these are not true (sub)games, because there is a terminal information set of the recipient across them, and this is potentially relevant for psychological reasons.

[8] The assumption of two heterogeneous groups is introduced for expositional simplicity. The more complex alternative would be to consider bidimensional types.

[9] The label D refers to the dictator, the agent who decides how to allocate the payoff, while R refers to the recipient, the passive player who receives the allocation. These labels are used as subscripts to distinguish the payoffs of the two players.

where $[y]^+ = \max\{0, y\}$, α represents the recipient's first-order probabilistic belief about the variables that determine their payoff, and parameter $\theta_g \geq 0$ gauges the dictator's guilt sensitivity, i.e., the individual aversion to let down (disappoint) the recipient. Note, as is now standard in psychological game theory, the dictator's utility depends on the recipient's belief, which is unknown to the dictator, who has to consult their second-order belief β to determine the subjectively optimal choice.[10]

For n-dictators, the indifference condition $x - \theta_n = M/2$ yields threshold

$$\hat{\theta}_n(x) = x - \frac{M}{2}, \tag{3}$$

which is the parameter value above (below) which a dictator chooses to share (not share). Note that $\hat{\theta}_n(x_H) - \hat{\theta}_n(x_L) = x_H - x_L > 0$. As long as interval $(x_L, x_H]$ contains a positive mass of n-types, increasing the endowment from x_L to x_H strictly decreases the sharing rate (which moves in the opposite direction of the threshold) in the subpopulation of n-dictators. This holds independently of agents' beliefs and chance probability p.

In our design, $\hat{\theta}_n(x_H) = 40 - 20 > \hat{\theta}_n(x_L) = 24 - 20$; thus, *for norm-driven dictators the sharing rate is higher when the material opportunity cost of sharing, parameterized by x, is lower.* We will refer to this as the **social-norm channel**.

Next we analyze the subpopulation of g-dictators. We refer to the effects of exogenous variations on behavior of g-dictators as the **guilt-aversion channel.** The effect of guilt aversion is isolated by looking at how second-order beliefs β affect behavior for each initial endowment x. However, we also have to analyze the effect of a change in endowment. Since it is commonly known that the realized endowment x is observed only by the dictator, recipients' first-order beliefs (FOBs) cannot depend on it, and it makes sense to assume that the dictators' second-order beliefs (SOBs) are independent of x as well. Accordingly, our analysis of how dictators' choices vary with x *holds second-order beliefs β constant.* We are going to show that, unlike n-dictators, *for g-dictators the sharing rate is increasing in x.* Here we prove this under a simplifying assumption, but the result is general (see

[10] See Battigalli and Dufwenberg (2022) for a comprehensive review and discussion of psychological game theory.

Battigalli *et al.* 2025).

Let $\bar{m}(\beta) := \mathrm{E}_{\beta}(\mathrm{E}_{\boldsymbol{\alpha}}(\mathbf{m}_R))$ denote the dictator's expectation of how much the recipient expects to get according to their SOB β. With this, under the simplifying assumption that *the dictator is certain that* $\mathrm{E}_{\boldsymbol{\alpha}}(\mathbf{m}_R) > M - x$ (i.e., that keeping induces positive disappointment), the expected utility of keeping is $x - \theta_g[\bar{m}(\beta) - (M - x)]$, with $\bar{m}(\beta) > (M - x)$. Since sharing cannot disappoint, its utility is simply $M/2$. Thus, the indifference condition $\theta_g[\bar{m}(\beta) - (M - x)] = x - M/2$ yields threshold

$$\hat{\theta}_g(\beta, x) = \frac{x - \frac{M}{2}}{\bar{m}(\beta) - (M - x)} = \frac{x - \frac{M}{2}}{x - (M - \bar{m}(\beta))}. \tag{4}$$

The equation shows that if the dictator's second-order expectation $\bar{m}(\beta)$ is low, i.e., close to $(M - x)$, the disappointment that the dictator expects to inflict is low and guilt sensitivity must be high to induce the dictator to share. Thus, *the sharing rate* (which moves in the opposite direction of the threshold) *is increasing in* $\bar{m}(\beta)$. Furthermore, $\lim_{\bar{m}(\beta) \searrow (M-x)} \hat{\theta}_g(x, \beta) = +\infty$.

In (4), the numerator represents the material (opportunity) cost of sharing, while the denominator is proportional to the expected psychological cost of keeping (i.e., of disappointing the recipient). An increase in x raises both costs. Thus, *these two effects partially cancel out.* However, unlike $\hat{\theta}_n(x)$, threshold $\hat{\theta}_g(\beta, x)$ is *decreasing* in x. To see this, note that $\bar{m}(\beta) \leq M/2$, because the recipients' most optimistic expectation is that the dictator shares with probability 1 independently of x, i.e., that they certainly get $M/2$. Thus, $M - \bar{m}(\beta) \geq M/2$, making the numerator in the right-hand side of (4) weakly larger than the denominator. This implies that an increase in x increases the numerator proportionally less than the denominator, and the fraction decreases.

Battigalli *et al.* (2025) show that the same conclusion holds also if the dictator is *not* certain to disappoint the recipient, and under more general assumptions about the psychological utility function of guilt-driven agents.

Observation 1 (Belief boundaries) *The lowest value of second-order expectation* $\bar{m}(\beta)$ *is attained when, according to SOB* β*, recipients are certain that dictators never share, which yields random recipient's payoff* $M - \mathbf{x}$*. The highest value*

of $\bar{m}(\beta)$ *is attained when, according to* β*, recipients are certain that dictators always share, which yields payoff* $M/2$ *with certainty. In our design, this implies treatment-dependent bounds on* $\bar{m}(\beta)$*. Letting* $\bar{m}(\beta|t)$ *denote the value of* $\bar{m}(\beta)$ *in treatment* $t \in \{t_1, t_2\}$*, we have* $\bar{m}(\beta|t_1) \in [12, 20]$ *and* $\bar{m}(\beta|t_2) \in [4, 20]$*. Note that the lower bound is not zero and differs across treatments because the probability of playing the* 24*-Game, in which* $M - x = 16$*, is not zero.*

As explained above, the social-norm channel and the guilt channel make sharing behavior react in opposite ways to an increase in dictators' endowment, with the former implying less sharing due to the higher opportunity cost. However, the guilt channel is the net result of two effects that partially cancel out. Therefore, it is plausible to assume that, if the fraction of dictators affected by the social-norm motivation is not much smaller than the complementary fraction of dictators affected by guilt aversion, the aggregate effect of an increase in dictators' endowment is to decrease sharing:

Auxiliary Hypothesis. Under plausible assumptions about the distribution of types, we expect that on average and keeping beliefs fixed, *the social-norm channel prevails, and the sharing rate is lower when the dictators' endowment is high.*

In the Appendix we provide conditions on the primitives of the model implying the prevalence of the social-norm channel.

Observation 2 *In each treatment of our design, beliefs* $\bar{m}(\beta)$ *are fixed, but the threshold for guilt aversion (4) varies depending on the endowment observed by the dictator. That is*

$$\begin{aligned} \hat{\theta}_g(\beta, x_H) &= \frac{40-20}{[40-(40-\bar{m}(\beta))]^+} = \frac{20}{\bar{m}(\beta)} \\ \hat{\theta}_g(\beta, x_L) &= \frac{24-20}{[24-(40-\bar{m}(\beta))]^+} = \frac{4}{[\bar{m}(\beta)-16]^+}, \end{aligned}$$

with $\hat{\theta}_g(\beta, x_L) = +\infty$ *if* $[\bar{m}(\beta) - 16]^+ = 0$*, i.e., if* $\bar{m}(\beta) \leq 16$*. The effects of guilt aversion are always triggered when the dictator is in the high-endowment condition* (x_H)*, whereas in the low-endowment condition* (x_L)*, they will only be triggered if* $\bar{m}(\beta) > 16$*.*

From the above observation, we can state the following. In our design, beliefs $\bar{m}(\beta)$ are independent of the endowment within each treatment, but they vary across treatments due to differences in the distribution of the random endowment $\mathbf{x}$. If the social-norm channel is dominant, and this is internalized in beliefs, we should expect $\bar{m}(\beta|t_1) > \bar{m}(\beta|t_2)$, and guilt aversion will imply corresponding differences in behavior across treatments. In the 40-Game:

$$\hat{\theta}_g(\beta, x_H|t_1) = \frac{20}{\bar{m}(\beta|t_1)} < \frac{20}{\bar{m}(\beta|t_2)} = \hat{\theta}_g(\beta, x_H|t_2).$$

Similarly, in the 24-Game:

$$\hat{\theta}_g(\beta, x_L|t_1) = \frac{20}{[\bar{m}(\beta|t_1) - 16]^+} \leqslant \frac{20}{[\bar{m}(\beta|t_2) - 16]^+} = \hat{\theta}_g(\beta, x_L|t_2),$$

where the equality holds if $\bar{m}(\beta|t_1) \leq 16$. In particular, when the dictator plays the 24-Game, guilt aversion is not triggered by our exogenous variation in beliefs, unless dictators' SOBs are very high in the high-belief treatment t_1.

The inequalities above identify the guilt-aversion channel. Our analysis suggests that this channel should more strongly affect behavior in the high-endowment condition (40-Game), than in the low-endowment condition (24-Game), and in the latter it may have a negligible effect.

2.3 Experimental hypotheses

Based on the theoretical model, we derive several testable predictions. When the dictator's endowment (and thus the opportunity cost of sharing) increases, the two motivational channels move in opposite directions: the social-norm channel predicts a lower sharing rate, because complying with the equal-sharing norm becomes more costly, whereas the guilt-aversion channel predicts a (weakly) higher sharing rate, because choosing Keep is expected to inflict more disappointment on the recipient. Under plausible assumptions on the distribution of types, we expect the social-norm channel effect to dominate, so that higher sharing costs reduce average sharing rates. Moreover, a higher probability of incurring large sharing costs should lower both first-order beliefs (FOBs) and second-order beliefs (SOBs).

If guilt aversion is present, this probability should also affect sharing through the guilt-aversion channel.

Our hypothesis is thus that recipients believe dictators are more likely to share when the opportunity cost of sharing is lower. Consequently, dictators' second-order beliefs (SOBs) are expected to be higher when the low-cost/endowment condition is more likely. Although recipients do not observe the actual realization of the endowment, their first-order beliefs (FOBs) depend on the treatment-probability that the low-cost condition obtains. Since this information structure is common knowledge, dictators internalize recipients' reasoning when forming their own expectations.

Under these hypotheses, both FOBs and SOBs are higher when the low-cost condition is more likely, leading to higher expected sharing, as shown in the theoretical section. Because our intended exogenous variation in beliefs relies on the Auxiliary Assumption of the previous subsection and its internalization in FOBs and SOBs, it is important to test whether it actually works. Specifically, we examine:

$$H1a \quad : \quad \text{fob}(t_1) > \text{fob}(t_2) \tag{5}$$

$$H1b \quad : \quad \text{sob}(t_1) > \text{sob}(t_2), \tag{6}$$

where $\text{fob}(t)$ and $\text{sob}(t)$ respectively denote the probability with which recipients expect dictators to share and the dictators' expectation of this probability in treatment $t \in \{t_1, t_2\}$.

The verification of (5) and (6) would support the existence of a predominant within-treatment social-norm channel. It is worth noting that this constitutes a sufficient condition for establishing the existence of such a channel, and a necessary condition for asserting its within-treatment prevalence over the guilt-aversion channel, since (5) and (6) follow from such prevalence (Auxiliary Assumption) and its internalization in first and second-order beliefs.

Our theoretical model also provides conditions for the activation of the guilt-aversion channel. Assuming that (6) holds, the treatment variation offers evidence

in support of the guilt channel if the following condition is satisfied:

$$H2a: \ \text{share}(t_1|x_H) > \text{share}(t_2|x_H), \tag{7}$$

where share$(t|x)$ denotes the average sharing rate observed in treatment t when the dictator is given by Chance endowment x. Hypothesis (7) implies that, under the guilt-aversion channel, a higher sharing rate is expected when the dictator holds higher beliefs about the recipient's expectations, conditional on a high endowment.

This test isolates the effect of higher second-order beliefs (SOBs) on sharing behavior, holding the endowment constant. If, given the high-endowment condition, sharing rates were consistently higher in the high-belief treatment compared to the low-belief one, this would support the presence of the guilt-aversion channel. In particular, it would indicate that dictators are more likely to share when they believe that recipients strongly expect them to do so and expect to inflict high disappointment by keeping.

We can further examine the treatment effect when the dictator is playing the 24-Game:

$$H2b: \ \text{share}(t_1|x_L) \geqslant \text{share}(t_2|x_L). \tag{8}$$

In this case, although we introduce an exogenous variation in sharing beliefs, this constitutes a sufficient — but not necessary — condition to provide evidence in support of guilt aversion. Based on our theoretical reasoning, we do not expect the 24-Game to activate guilt aversion effects, unless the expected sharing rate is very high. This is because keeping (i.e., allocating 16 tokens to the recipient) is unlikely to cause disappointment, as the recipient's expected payoff is likely to fall below that threshold.

Hypotheses (9) and (10) below provide evidence consistent with the presence of the social-norm channel. The estimation of its effects relies on comparing sharing rates under the low and high-endowment conditions within each treatment. Formally:

$$H3a \ : \ \text{share}(t_1|x_L) > \text{share}(t_1|x_H), \tag{9}$$

$$H3b \ : \ \text{share}(t_2|x_L) > \text{share}(t_2|x_H). \tag{10}$$

The verification of these inequalities provides a sufficient (though not necessary) condition for the influence of the social-norm channel, given that — within treatments — the guilt-averse channel affects behavior in the opposite direction.

Furthermore, we also expect the following diff-in-diff inequality:

$$H4 : \text{share}(t_2|x_L) - \text{share}(t_2|x_H) > \text{share}(t_1|x_L) - \text{share}(t_1|x_H).$$

This should hold because we expect the countervailing guilt-aversion channel to be stronger in the high-beliefs treatment t_1 than in the low-beliefs treatment t_2, and in the latter the guilt-aversion channel may even be inactive.

3 Experimental procedure

The experiment was conducted at Tor Vergata University (Rome) in May 2022 and involved 512 undergraduate students across 16 sessions, recruited through an online recruitment system. The entire experiment was computerized and implemented using the *z-Tree* software (Fischbacher, 2007). Upon arrival at the laboratory, participants were randomly assigned to individual and isolated computer terminals. Two research assistants distributed the instructions and ensured that participants followed the procedures correctly. Before beginning the main tasks, all subjects completed a brief comprehension questionnaire to verify their understanding of the instructions.

Each session consisted of 8 rounds (with perfect stranger matching) in a within-subject design. Our design is built on the simple game described in Figure 1. We always let $M = 40$, $x_L = 24$, and $x_H = 40$. A high opportunity cost of sharing characterizes the 40-Game, i.e., the equal payoff distribution costs 20 tokens to the dictator in the 40-Game and only 4 tokens in the 24-Game.

Each round implemented the following sequence of a five-stage procedure.

1. Role assignment. Player roles A and B are randomly assigned, and pairs are formed.

2. Probability of distribution (treatment). Participants are informed about the probability p that the round will be a 24-Game (i.e., that the dictator's

endowment will be $x_L = 24$). In each round, p can be either 0.75 (t_1) or 0.25 (t_2).

3. Belief elicitation. Recipients' FOBs and dictators' SOBs are elicited before the endowment realization. This stage has two parts: (a) first-order beliefs: each recipient predicts the dictator's action; (b) second-order beliefs: each dictator predicts the recipient's first-order belief.

4. Endowment assignment. Endowments are randomly assigned according to p. Only dictators observe the realized endowment (asymmetric information).

5. Dictator's action. Dictators choose whether to share equally (20–20) or to keep the randomly determined status quo.

At the end of each session, one round was randomly selected for payment based on the allocation decision. Belief elicitation was instead incentivized in all remaining rounds.[11] All the game payoffs were described in terms of "tokens," with 1 token = 0.50 euros. In addition, each subject received a fixed show-up fee of 3 tokens.

4 Experimental results

Our results are displayed in Table 1, which reports the elicited average FOBs and SOBs (about how likely the dictator is to share) and sharing rates. Experimental outcomes are structured by treatment (t_1 and t_2; rows) and endowment (24- vs. 40-Game; columns). The table also reports the standard deviations below the sharing rates. Overall, average sharing rates vary between 35% and 46% across

[11] The belief-elicitation procedure is adapted from Vanberg (2008) and proceeds as follows (see the Appendix for details). First, each recipient states a prediction about the dictator's action. Second, each dictator reports her prediction of the recipient's first-order belief. Beliefs were incentivized in rounds that were not used to determine monetary payoffs. Similar elicitation schemes are widely used in the experimental literature (e.g., Di Bartolomeo *et al.*, 2019; Ciccarone *et al.*, 2020; Di Bartolomeo *et al.*, 2023a, 2023b, 2024; Charness *et al.*, 2024). This procedure reduces incentives to hedge between actions and belief reports. However, as discussed by Cox *et al.* (2015), hedging concerns cannot be entirely ruled out when multiple incentivized tasks are combined.

treatments and endowments. Average FOBs and SOBs range between 0.35 and 0.49.

Table 1 – FOBs, SOBs, and sharing rates

	FOBs	SOBs			Sharing rate	
	All	All	40-G.	24-G.	40-G.	24-G.
	(a)	(b)	(c)	(d)	(e)	(f)
(1) High prob. to	48%	48%	49%	48%	42%	46%
play 24-Game (t_1)	0.31	0.31	0.31	0.31	0.49	0.50
(2) Low prob. to	35%	36%	36%	36%	35%	41%
play 24-Game (t_2)	0.31	0.31	0.31	0.31	0.48	0.49

NOTE: The table reports averages and standard deviations of FOBs, SOBs, and sharing rates for each treatment and for each initial endowment observed by dictators (24 or 40). In columns (a) and (b), each cell contains 1024 observations, whereas in the remaining columns, each cell contains 512 observations.

The manipulation of the probability of playing the low–endowment game (24-Game) works as intended ($H1$): when the 24-Game is more likely to be played, both FOBs and SOBs that the dictator will share are substantially higher than when it is less likely. On average, subjects are less confident that dictators will share when the material opportunity cost of sharing is more likely to be high. As explained, this result is consistent with the existence of a (predominant) social-norm channel.

In columns (a) and (b) of Table 1, we respectively observe 48% vs. 35% for FOBs and 48% vs. 36% for SOBs (Wilcoxon signed-rank test, $Z = 3.52$, $p = 0.000$).[12] All statistics reported in the paper are based on Wilcoxon signed-rank tests comparing averages at the session level. Our data are independent

[12] We also compare SOBs between dictators who later play the 40- and 24-Game (columns (c) and (d)). As expected, these are unaffected by the randomly selected subgame, since expectations are elicited before dictators learn which subgame they will play (tests omitted).

across sessions but not across individuals; the Wilcoxon tests account for this data structure.

The exogenous variation of probability p of the low endowment allows us to compare, ceteris paribus, two different levels of SOBs and thereby test for the guilt aversion channel. As implied by the theoretical model, we expect this effect to be larger when keeping yields a payoff of zero to the recipient ($H2a$). By contrast, in our model the effects of guilt aversion are small or negligible when keeping grants 16 tokens to the recipient, because in this case it is unlikely that such a choice will generate (and be expected to generate) substantial disappointment in the recipient ($H2b$).

Our empirical strategy builds on the exogenous variation in beliefs induced by our experimental design: we isolate the guilt-aversion channel and compare sharing rates within columns (e) and (f), which correspond to hypotheses $H2a$ and $H2b$, respectively. Dictators with higher SOBs appear more likely to share, both when facing high and low endowments. Consistent with the theoretical model, the effect is stronger and statistically significant in the high-endowment condition. When keeping leaves the recipient with zero ($H2a$), we observe a 7 p.p. difference in sharing rates between the high- and low-belief treatment (42% vs. 35%; $Z = 1.81$, one-sided $p = 0.035$). In the low-endowment condition ($H2b$), the difference in sharing rates is smaller and only marginally significant: 46% vs. 41% ($Z = 1.50$, one-sided $p = 0.066$).

Summing up, $H2a$ provides clear evidence of a guilt-aversion effect, since in this case keeping is expected to substantially disappoint the recipient. By contrast, $H2b$ does not reach conventional significance levels, which is consistent with the theoretical prediction that the guilt-aversion channel should be weak or negligible when keeping leaves the recipient with 16 tokens and is therefore unlikely to generate disappointment.

Next, we turn to non–belief-dependent motivations, i.e., the social-norm channel ($H3$). Our exogenous variation in the endowment (hence, in the opportunity cost of sharing) allows us to observe its effect by comparing sharing rates within rows, i.e., holding beliefs fixed within each treatment. As argued above, dictators facing a lower material cost of sharing should be more likely to choose *share*—both when SOBs are high and when they are low—because complying with the equal-

sharing norm is less demanding. However, changes in the opportunity cost also activate the (non-predominant) guilt-aversion channel, which may partially offset the social-norm one. Since the guilt-aversion channel is stronger when expectations are high, we expect the net impact of a fall in the opportunity cost of sharing to be smaller in the high-belief treatment t_1 than in the low-belief treatment t_2 (diff-in-diff hypothesis $H4$).

Table 1 shows that, in both treatments, the sharing rate is higher when the initial endowment is lower: it rises from 42% to 46% in t_1 and from 35% to 41% in t_2. Taken together with the results for $H1$, these directional patterns support the presence of a (predominant) social-norm channel. Moreover, the effect is larger (6 p.p. vs. 4 p.p.) when the probability of being in the 24-Game group is smaller, as predicted by the theory: in this treatment (t_2), the mitigating effect of the countervailing guilt-aversion channel is lower. However, the differences are not statistically significant. In the high-belief treatment (t_1) we observe 46% vs. 42%: $Z = 0.10$, $p = 0.541$ ($H3a$); in the low-belief treatment (t_2), we observe 41% vs. 35%: $Z = 1.14$, $p = 0.127$ ($H3b$). It follows that our results are directionally consistent with our diff-in-diff hypothesis ($H4$), but the difference is not statistically significant ($Z = 0.90$, $p = 0.365$).

Our experimental design generates exogenous variation in second-order beliefs (SOBs), effectively creating a high- versus low-belief environment. Beyond this treatment-level shift, we also elicit each dictator's individual SOB. Exploiting this richer, individual-level measure allows us to move beyond session averages and examine sharing behavior at a finer level of detail. The corresponding results are reported in Table 2 and illustrated in Figure 2.

Table 2 reports the estimates from a linear regression where the dependent variable is a binary measure of the dictator's action—equal to 100 if the high-share option was chosen, and 0 otherwise. We prefer to estimate linear probability models rather than probit or logit because, in our setting, sharing rates fall in a relatively narrow mid-range interval (35–46%), so the linear specification is a good local approximation to nonlinear models and its coefficients can be read directly as percentage-point changes in the probability of choosing share.[13] The key

[13]Coefficients can be interpreted as marginal effects—percentage point changes in the probability of choosing the share option. SOBs and the round variable are standardized; the reference

explanatory variables are the dictator's SOB, the endowment condition (24 or 40), and their interaction, with an additional control for the round in which the decision was made. Figure 2 complements the regression analysis by visually presenting the marginal effects of SOBs on sharing behavior, conditional on endowment.

SOBs and the endowment map naturally into our two mechanisms. A higher SOBs means the dictator believes the recipient expects sharing with higher probability, so guilt aversion predicts a positive coefficient on SOBs. The endowment dummy equals 1 when the dictator's endowment is high (40), i.e., when sharing is more costly (the opportunity cost of choosing share is larger). Cost-based motives—including compliance with an equal-split norm when it is less demanding—therefore predict a negative coefficient on the endowment dummy.

A key caveat is that SOBs are not a direct measure of the dictator's expectation of the recipient's disappointment. They measure the dictator's expectation of the probability assigned by the recipient to the sharing action, but whether the Keep action is disappointing depends on how much the recipient obtains, which in turn depends on the endowment observed only by the dictator: for fixed FOBs, the recipient is more disappointed when the endowment that the dictator keeps is high.

For this reason, the interaction term should be read as capturing whether the belief effect is amplified when the endowment is high. A positive interaction is consistent with the idea that, at higher endowments, keeping is more likely to fall short of what the recipient expects, which strengthens the guilt motive and can partially offset the larger opportunity cost of sharing.

Table 2 – Sharing behavior regression

	coef.	s.e. rob.	t	p-value	95% CI	
	(a)	(b)	(c)	(d)	(e)	(f)
(1) SOBs	5.69	1.85	3.08	0.002	2.06	9.33
(2) endowment	−5.68	2.21	−2.57	0.010	−10.01	−1.35
(3) (1)×(2)	1.77	2.26	0.78	0.433	−2.67	6.21
(4) round	−1.55	0.93	−1.66	0.098	−3.38	0.28
(5) constant	44.00	1.10	39.87	0.000	41.83	46.17

category corresponds to average belief (42.3%) and average round (4.5).

NOTE: The table reports the results of a linear regression where the dependent variable is a binary measure of the dictator's decision, equal to 100 if the share option was chosen, and 0 otherwise. The key independent variables are the dictator's SOB, the endowment condition (24 or 40), and their interaction, along with a control for the round in which the decision was made. The specification allows for a direct interpretation of the coefficients as marginal effects, i.e., p.p. changes in the probability of choosing the high-share option. SOBs and the round variable have been standardized so that the constant can be interpreted as the average probability when the endowment is 24 (average SOBs 42.3% and round 4.5).

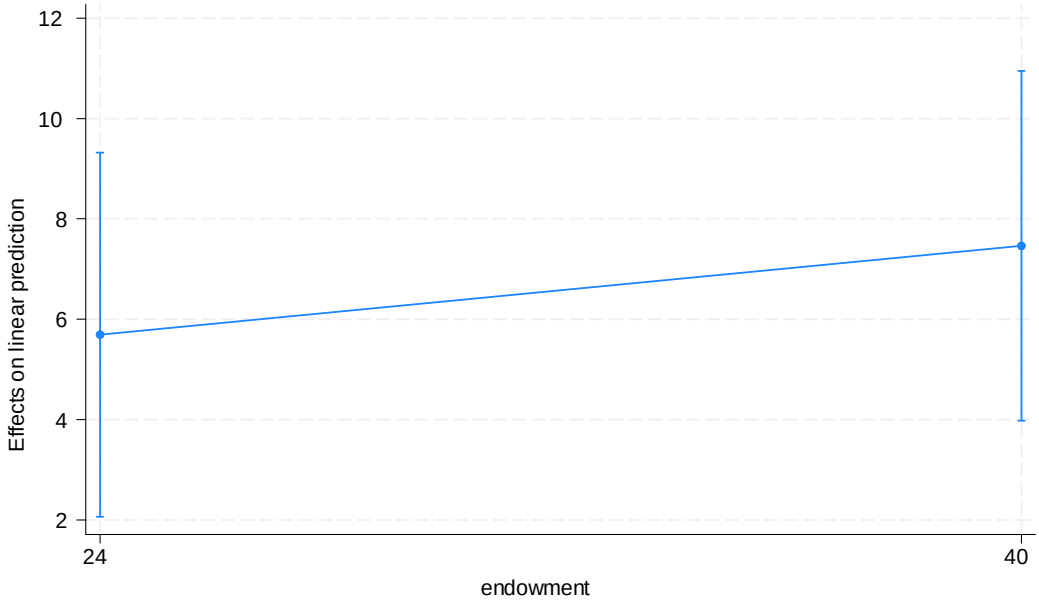

Figure 2 – Marginal effect of SOBs (95% CI).

NOTE: The figure displays the estimated marginal effects of SOBs on the probability of choosing the share option, conditional on the

endowment level (24 or 40), with 95% confidence intervals.

Table 2 shows that higher SOBs are associated with a significantly greater probability of sharing: a one-standard-deviation increase in SOBs raises the likelihood of choosing the share option by about 5.7 p.p. ($p = 0.002$), in line with guilt-aversion predictions.

The above result implies that dictators with the highest SOBs have a 10.4 p.p. higher probability of choosing the share option compared to those with beliefs around the average, while those with the lowest SOBs are 7.6 p.p. less likely to do so.[14] For instance, holding all other variables at their mean values, in the low-cost environment, the estimated probability of sharing is 36.4% when the SOBs are at their lowest value, whereas it is 54.4% when the SOBs are at their highest value. By construction, when the SOBs are at their mean value, the estimated probability is 44%, i.e., the value of the estimated constant term.

The negative and significant coefficient on the high-endowment dummy indicates that a higher material opportunity cost of sharing reduces the propensity to share by 5.68 p.p., capturing the social-norm channel. That is, holding all other variables at their mean values (SOBs and round), the expected sharing rate is 44.0% for the low-cost endowment, whereas it is 38.3% for the high-cost endowment.

Consistently with the presence of the guilt-aversion channel, the interaction term between SOBs and endowment is positive, although not statistically significant, suggesting that the influence of beliefs tends to become stronger when keeping gives less to the recipient.

Finally, there is a negative time trend, consistent with some learning or adaptation over rounds.

Our theoretical organization of the data is further illustrated by Figure 2, which shows that the effect of SOBs is positive in both endowment conditions, and slightly larger when dictators' endowment is 40. While the difference is not statistically significant, the observed pattern directionally aligns with the experimental hypotheses.

[14] The percentages are obtained by multiplying the marginal effect by the standardized belief values of the highest and lowest SOB groups.

In sum, the data indicate that variation in the endowment (i.e., an increase in the opportunity cost) is associated with a decrease of about 5 p.p. in the probability of choosing sharing. Moreover, we find evidence of a guilt-aversion effect: the impact can be as large as 18 p.p. when comparing dictators with lowest SOBs to those with highest SOBs. The estimates suggest that guilt aversion is more effective at promoting sharing when the endowment (and thus the opportunity cost) is higher. Nevertheless, the corresponding interaction effect is not statistically significant.

5 Additional results: Gender effects

We next ask whether these mechanisms operate differently across genders. Motivated by evidence from experiments in similar settings that reveal interesting gender patterns,[15] we re-estimate our baseline specification separately for female and male dictators. This analysis should be viewed as exploratory and mainly hypothesis-generating, aimed at shedding light on potential heterogeneity in how guilt aversion and social norms shape sharing behavior for females and males.

Our results are reported in Tables 3 and 4, which present the same regression specification as in Table 2, estimated, respectively, for female and male dictators.

According to the results of Table 3 (females), SOBs have a positive but only weakly significant effect on sharing ($p \approx 0.10$), while the main effect of the high-endowment dummy is not significant. However, the interaction between SOBs and endowment is positive and statistically significant, indicating that beliefs are especially important for females when the endowment is high. In other words, females' sharing decisions appear to react more strongly to SOBs precisely when keeping should be expected to inflict more disappointment. The round effect shows a negative trend over time, possibly reflecting learning effects.

[15] See, among others, Charness *et al.* (2024) and Di Bartolomeo *et al.* (2024).

Table 3 – Sharing behavior regression: females

	coef.	s.e. rob.	t	p-value	95% CI	
	(a)	(b)	(c)	(d)	(e)	(f)
(1) SOBs	4.32	2.60	1.66	0.098	−0.81	9.44
(2) endowment	−2.81	3.15	−0.89	0.374	−9.03	3.40
(3) (1)×(2)	6.68	3.09	2.17	0.031	0.60	12.77
(4) round	−2.97	1.22	−2.43	0.016	−5.38	−0.56
(5) constant	43.40	1.57	27.70	0.000	40.32	46.17

Table 4 shows the regression results for male dictators. In this subsample, SOBs are strongly and significantly correlated with sharing: a one-standard-deviation increase in SOBs raises the probability of sharing by about 6.7 p.p. ($p = 0.010$). At the same time, males react more sharply to the endowment: the high-endowment dummy is negative and significant, implying that higher opportunity costs of sharing substantially reduce it. The interaction between SOBs and endowment is not significant and slightly negative, suggesting that, for males, beliefs matter in both endowment conditions, but are not disproportionately amplified when endowment is high. Overall, male behavior reflects a combination of guilt aversion and sensitivity to material costs, with clearer evidence of a social-norm effect operating through the endowment term.

Table 4 – Sharing behavior regression: males

	coef.	s.e. rob.	t	p-value	95% CI	
	(a)	(b)	(c)	(d)	(e)	(f)
(1) SOBs	6.74	2.58	2.61	0.010	1.66	11.83
(2) endowment	−8.22	3.06	−2.69	0.008	−14.24	−2.21
(3) (1)×(2)	−2.01	3.21	−0.63	0.531	−8.33	4.30
(4) round	−0.34	1.38	−0.25	0.806	−3.05	2.37
(5) constant	44.51	1.54	28.96	0.000	41.48	47.53

Overall, females' behavior looks more consistent with the guilt-aversion channel (beliefs "bite" especially when endowment is high), whereas males display a cleaner combination of guilt aversion (SOBs) and social norm sensitivity (endowment) with

little evidence that endowment systematically amplifies the belief channel. Females also show a significant negative round trend, indicative of learning/adaptation over time. The round effect is negligible and not significant for males.

Given that this split-sample exercise is exploratory, the contrast should be interpreted as suggestive heterogeneity rather than definitive proof of gender differences.

6 Conclusions

This paper studies how prosocial behavior is shaped by the interaction of two mechanisms emphasized in the social-preferences literature. The first is a social-norm channel, in which the dictators' endowment (hence, the material opportunity cost of sharing) affects the sharing rate, and therefore, how much sharing is expected according to the probability of the low-endowment condition, the treatment variable of our design. The second is a guilt-aversion channel, in which dictators' choices respond to second-order beliefs—their beliefs about what the recipient expects—because deviating from those expectations generates psychological costs.

A theoretical model motivates our experimental design and yields testable predictions about how sharing should vary with (i) dictators' endowment and (ii) second-order beliefs, as well as how the two may interact. Guided by this framework, we implement a modified dictator game in which the dictator's endowment is random and realized after belief elicitation, and the recipient does not observe the realized endowment. We manipulate the publicly known probability of facing a low- versus high-endowment game to shift first- and second-order beliefs, while the realized endowment generates variation in the material cost of sharing.

Our results show, first, that the belief manipulation works as intended: when the low-endowment game is more likely, both first-order and second-order beliefs about sharing increase. Turning to behavior, sharing is positively related to second-order beliefs, consistent with guilt aversion: dictators who think the recipient expects them to share are more likely to do so. At the same time, within each treatment, higher endowments reduce sharing, indicating that material incentives and the norm-driven channel are strong enough to overcome the countervailing guilt-channel given fixed beliefs. Overall, the evidence supports a joint role for the

two channels: the cost environment shapes expectations, and those expectations in turn predict sharing.

We also explore heterogeneity by gender. For female dictators, the association between second-order beliefs and sharing is stronger when the endowment is high, while for male dictators beliefs and costs display clearer separate effects with a weaker interaction. These patterns are exploratory and should be interpreted as hypothesis-generating.

The study has limitations: it uses a laboratory subject pool and a simple two-endowment structure in a one-shot sharing environment. Future work could extend the framework to richer strategic settings, alternative norm structures, and different information regimes, and further investigate how individual characteristics interact with guilt aversion and norm adherence. Despite these limitations, the paper provides theory-guided evidence that prosocial behavior reflects an interplay between social norms, economic incentives, and belief-dependent motivations.

Appendix

A1. Belief elicitation

Elicitation of first-order beliefs. Recipients make their guess by ticking one of the five points in Table A (Vanberg, 2008). Beliefs are then re-scaled to 1, 0.75, 0.50, 0.25, and 0. Thus, the numbers in Table 1 represent averages of recipients' re-scaled responses. The payoffs correspond to a quadratic scoring rule for probability values 85%, 68%, 50%, 32%, and 15% (see Vanberg, 2008, p. 1472).

Table A – Incentives for the elicitation of first-order beliefs (tokens)

A will	choose *Share*			choose *Keep*	
	Certainly	Probably	Unsure	Probably	Certainly
Please tick your guess	O	O	O	O	O
Your earnings if A chooses					
Share	0.65	0.60	0.50	0.35	0.15
Keep	0.15	0.35	0.50	0.60	0.65

Elicitation of second-order beliefs. After recipients submit their first-order belief, dictators are asked to guess which of the five points in Table A was selected by their counterpart. Correct guesses are paid 0.50 tokens.

A2. A sufficient condition for the Auxiliary Hypothesis

We present a condition about the primitives of the model that implies the Auxiliary Hypothesis, i.e., that — keeping beliefs fixed — the norm-channel for changes in endowments prevails, and thus the sharing rate is strictly higher with low endowment x_L than with high endowment x_H.

Let F_n (resp. F_g) denote the cdf of $\boldsymbol{\theta}_n$ (resp. $\boldsymbol{\theta}_g$) in the subpopulation of norm-driven (resp. guilt-driven) dictators, and recall that q_n denotes the fraction of norm-driven dictators. As will be clear from the proof, for the sufficient condition we are going to derive, it is without loss of generality to assume that all guilt-driven dictators have the same SOB β. With this, the sharing rate with endowment x and SOB β is

$$\Pr(\text{Share}; x, \beta) = q_n \left(1 - F_n\left(\hat{\theta}_n(x)\right)\right) + (1 - q_n)\left(1 - F_g\left(\hat{\theta}_g(x, \beta)\right)\right). \quad (11)$$

Proposition 3 *The norm-channel prevails and the sharing rate is higher when the endowment is low if*

$$q_n \left(F_n\left(x_H - \frac{M}{2}\right) - F_n\left(x_L - \frac{M}{2}\right)\right) > (1 - q_n)(1 - F_g(1)).$$

Proof. Recall that the indifference threshold for a norm-driven dictator with endowment $x \in \{x_L, x_H\}$ is

$$\hat{\theta}_n(x) = x - \frac{M}{2}.$$

Given that in the parametrization of our experiment a dictator with high endowment is certain to disappoint the recipient, the indifference threshold for a guilt-driven dictator with *high endowment* is

$$\hat{\theta}_g(\beta, x_H) = \frac{x_H - \frac{M}{2}}{\bar{m}(\beta) - (M - x_H)} = \frac{M/2}{\bar{m}(\beta)} \geq 1,$$

because $x_H = M$ and $\bar{m}(\beta) := \mathrm{E}_\beta(\mathrm{E}_{\boldsymbol{\alpha}}(\mathbf{m}_R)) \in [0, M/2]$. Note that

$$\inf_\beta \hat{\theta}_g(\beta, x_H) = 1 \tag{12}$$

is attained when $\bar{m}(\beta) = M/2$, i.e., when the dictator is certain that the recipient is certain that they are going to share.

When the endowment is *low*, the expected disappointment is also low, and the dictator has little or no incentive to share. In particular,

$$\sup_\beta \hat{\theta}_g(\beta, x_L) = +\infty \tag{13}$$

is attained for every SOB β such that $\mathrm{P}_\beta(\mathrm{E}_{\boldsymbol{\alpha}}(\mathbf{m}_R) \leq M - x_L) = 1$.

Therefore,

$$\mathrm{Fr}(\text{Share}; x_L, \beta) > \mathrm{Fr}(\text{Share}; x_H, \beta)$$

if and only if

$$\begin{aligned} & q_n\left(1 - F_n\left(x_L - \frac{M}{2}\right)\right) + (1 - q_n)\left(1 - F_g\left(\hat{\theta}_g(x_L, \beta)\right)\right) \\ > \; & q_n\left(1 - F_n\left(x_H - \frac{M}{2}\right)\right) + (1 - q_n)\left(1 - F_g\left(\hat{\theta}_g(x_H, \beta)\right)\right). \end{aligned}$$

Since

$$\begin{aligned} & \mathrm{Fr}(\text{Share}; x_L, \beta) - \mathrm{Fr}(\text{Share}; x_H, \beta) \\ \geq \; & \inf_\beta \left(\mathrm{Fr}(\text{Share}; x_L, \beta) - \mathrm{Fr}(\text{Share}; x_H, \beta)\right) \\ \geq \; & \inf_\beta \mathrm{Fr}(\text{Share}; x_L, \beta) - \sup_\beta \mathrm{Fr}(\text{Share}; x_H, \beta), \end{aligned}$$

eq.s (11), (12), and (13) imply that sharing is higher under low endowment if

$$q_n\left(1 - F_n\left(x_L - \frac{M}{2}\right)\right) > q_n\left(1 - F_n\left(x_H - \frac{M}{2}\right)\right) + (1 - q_n)(1 - F_g(1)),$$

that is, if

$$q_n\left(F_n\left(x_H-\frac{M}{2}\right)-F_n\left(x_L-\frac{M}{2}\right)\right)>(1-q_n)\left(1-F_g\left(1\right)\right). \blacksquare$$

References

Attanasi, G., P. Battigalli, and E. Manzoni (2016), "Incomplete information models of guilt aversion in the trust game," *Management Science*, 62, 648-667.

Attanasi, G., P. Battigalli, E. Manzoni, and R. Nagel (2025), "Disclosure of belief-dependent preferences in a trust game," *Economic Theory*, 10.1007/s00199-025-01645-5.

Bartling, B., F. Engl, and R. A. Weber (2014), "Does willful ignorance deflect punishment? An experimental study," *European Economic Review*, 70: 512-524.

Battigalli, P., and M. Dufwenberg (2007), "Guilt in games," *American Economic Review*, 7: 170-176.

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

Battigalli P., R. Corrao, and M. Dufwenberg (2019): "Incorporating belief-dependent motivation in games," *Journal of Economic Behavior and Organization*, 167, 185-218.

Battigalli, P., G. Di Bartolomeo, and S. Papa (2025), "Guilt aversion and inequality in dictator games," *Economics Letters*, 247, February.

Benabou, R., and J. Tirole (2006), "Incentives and prosocial behavior," *American Economic Review*, 96(5): 1652-1678.

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

Charness, G., G. Di Bartolomeo, and S. Papa (2024), "A stranger in a strange land: Promises and identity," *Games and Economic Behavior,* 144, 13-28.

Chowdhury, S., M. and J. Y. Jeon (2014), "Impure altruism or inequality aversion? An experimental investigation based on income effects," *Journal of Public Economics*, 118(C): 143-150.

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

Cox, J. C. (2004). "How to identify trust and reciprocity," *Games and Economic Behavior*, 46: 260-281.

Cox, J. C., V. Sadiraj, and U. Schmidt (2015), "Paradoxes and mechanisms for choice under risk." *Experimental Economics*, 18, 215-250.

d'Adda, G., M. Dufwenberg, F. Passarelli and G. Tabellini (2020), Social norms with private values: Theory and experiments." *Games and Economic Behavior*, 124: 288-304.

Danilov, A., K. Khalmetski, and D. Sliwka (2021), "Descriptive norms and guilt aversion," *Journal of Economic Behavior and Organization*, 191: 293–311.

Di Bartolomeo, G., and S. Papa (2016), "Trust and reciprocity: Extensions and robustness of triadic design," *Experimental Economics*, 19: 100-115.

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

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

Di Bartolomeo, G., M. Dufwenberg, and S. Papa (2023b), "The sound of silence: A license to be selfish," *Economics Letters,* 182, 68-70.

Di Bartolomeo, G., M. Dufwenberg, S. Papa, and L. Razzolini (2024), "Guilt aversion and moral commitment: Eve versus Adam," *Journal of the Economic Science Association*, 1-17.

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

Fehr, E., and K. M. Schmidt (1999), "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 114(3): 817-868,

Fehr, E., and K. M. Schmidt (2010), "On inequity aversion: A reply to Binmore and Shaked," *Journal of Economic Behavior and Organization*, 73: 101-108.

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

Gächter, S., L. Gerhards, and D. Nosenzo (2017), "The importance of peers for compliance with norms of fair sharing," *European Economic Review*, 97: 72-86.

Khalmetski, K. (2016), "Testing guilt aversion with 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.

Korenok, O., E. L. Millner, L. Razzolini (2012), "Are dictators averse to inequality?," *Journal of Economic Behavior and Organization*, 82(2-3): 543-547.

Korenok, O., E. L. Millner, L. Razzolini (2014), "Taking, giving, and impure altruism in dictator games," *Experimental Economics*, 17(3): 488-500.

Krupka, E. L., and R. A. Weber (2013), "Identifying social norms using coordination games: Why does dictator game sharing vary?," *Journal of the European Economic Association*, 11(3): 495-524.

Murnighan, J. K., and J. M. Oesch, M. Pillutla (2001), "Player types and self-impression management in dictator games: Two experiments," *Games and Economic Behavior*, 37: 388-414.

Nguyen, C. V. (2019), "The effect of inequality in stakes on sharing behavior: Evidence from an experimental study," *Economics Letters* 184, 108638.

Rothenhäusler, D., N. Schweizer, and N. Szech (2018), "Guilt in voting and public good games," *European Economic Review*, 101, 664-681.

Servátka, M. (2009), "Separating reputation, social influence, and identification effects in a dictator game," *European Economic Review*, 53: 197-209.

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

Tech Appendix

Instructions

Thank you for participating in this experiment. The experiment studies how people make decisions. If you have questions, please raise your hand. Please do not talk to other participants during the session. All decisions are anonymous.

Earnings

During the experiment you will earn *tokens*. At the end of the session, tokens will be converted into euros and paid in cash. In addition, everyone receives a fixed participation fee of **3 tokens**, independent of choices. The exchange rate is:

$$1 \text{ token} = 0.50 \text{ euros}.$$

General Structure

The experiment consists of **8 rounds**. In each round, you will be randomly paired with a different participant (perfect stranger matching). In each round, you will be randomly assigned one of two roles:

- **Participant A** (the decision maker)
- **Participant B**

Roles are randomly re-assigned in every round.

In each round, the total amount of tokens at stake is always **40**, but the *initial allocation* can be one of the following:

- **Endowment H:** A has 40 tokens and B has 0 tokens (denoted $[40, 0]$).
- **Endowment L:** A has 24 tokens and B has 16 tokens (denoted $[24, 16]$).

Sequence of Each Round

Each round consists of five phases.

Phase 1: Role Assignment and Matching The computer randomly assigns roles (A/B) and randomly forms pairs.

Phase 2: Probability Announcement Both participants are informed about the probability that the round will start from endowment $[40, 0]$ rather than $[24, 16]$. In a given round this probability is either:

- **75**% that the endowment is $[40, 0]$ and **25**% that it is $[24, 16]$, *or*
- **25**% that the endowment is $[40, 0]$ and **75**% that it is $[24, 16]$.

This probability is common knowledge in the round.

Phase 3: Belief Elicitation (Additional Rewards) Before the endowment is realized, both participants are asked to answer incentivized questions about expectations.

1. **First-order belief (B's belief).** Participant B is asked to predict whether A will choose *Share* or *Keep* in this round, using the following five-point scale:

   Certainly Share / Probably Share / Unsure / Probably Keep / Certainly Keep.

   This guess is paid according to the table below (tokens).

2. **Second-order belief (A's belief).** Participant A is asked to predict *which of the five options* participant B selected in the previous question. A correct guess is paid **0.50 tokens**.

Incentives for B's belief (tokens)

	Certainly Share	Probably Share	Unsure	Probably Keep	Certainly Keep
Earnings if A chooses Share	0.65	0.60	0.50	0.35	0.15
Earnings if A chooses Keep	0.15	0.35	0.50	0.60	0.65

Phase 4: Endowment Realization (Asymmetric Information) The computer draws the initial endowment according to the probability announced in Phase 2. **Only participant A observes** whether the realized endowment is $[40, 0]$ or $[24, 16]$. Participant B does **not** observe the realized endowment.

Phase 5: A's Decision and Information to B Participant A chooses between the following two actions:

- **Share:** the 40 tokens are split equally, so both receive $[20, 20]$.
- **Keep:** the initial endowment is implemented, i.e., A keeps the realized allocation ($[40, 0]$ or $[24, 16]$).

At the end of the round, participant B observes whether A chose *Share* or *Keep*, but does **not** learn which endowment A had.

Payoffs from A's Decision (tokens)

Realized endowment	If A chooses Share		If A chooses Keep	
$[40, 0]$	A: 20	B: 20	A: 40	B: 0
$[24, 16]$	A: 20	B: 20	A: 24	B: 16

Payment Rule

After all 8 rounds are completed, the computer randomly selects **one round** for payment of the **decision payoffs** (A's choice). Belief-elicitation earnings are paid for **all other rounds**, i.e., for every round *except* the one selected for decision payment. Finally, the participation fee of **3 tokens** is added.

Important Rules

- Do not communicate with other participants.
- Do not attempt to reveal your identity or participant code.
- Participants who violate these rules will be excluded and will not receive payments.