CBESS Discussion Paper 16-05R

University of East Anglia

cbess
Centre for Behavioural and Experimental Social Science

The Role of Morals in Three-Player Ultimatum Games

by Sandro Casal *
Francesco Fallucchi **
Simone Quercia***

*Department of Economics, University of Milan
** School of Economics, University of East Anglia
***Institute for Applied Microeconomics, University of Bonn

Abstract
We experimentally investigate the role of responders' moral concerns in three-player ultimatum bargaining. In our experiment, proposers can increase their share of the pie at the expenses of an NGO that conducts humanitarian aid in emergency areas. We find that responders are not willing to engage in 'immoral' transactions only when fully informed about proposers' behavior toward the NGO. Under complete information, their willingness to reject offers increases with the strength of the harm to the NGO. Moreover, the possibility to nullify the effects of the negative externality through rejection further increases their willingness to reject. We show that the latter result is better explained by a model of consequentialist moral concerns toward the NGO rather than deontological morality about own actions.

JEL classification codes
C72, C91, D6

Keywords
Three-player ultimatum game, moral reasoning, experiment.

Centre for Behavioural and Experimental Social Science
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/cbess

The Role of Morals in Three-Player Ultimatum Games

Sandro Casal[1], Francesco Fallucchi[2] and Simone Quercia[3]

May 10, 2017

Abstract

We experimentally investigate the role of responders' moral concerns in three-player ultimatum bargaining. In our experiment, proposers can increase their share of the pie at the expenses of an NGO that conducts humanitarian aid in emergency areas. We find that responders are not willing to engage in 'immoral' transactions only when fully informed about proposers' behavior toward the NGO. Under complete information, their willingness to reject offers increases with the strength of the harm to the NGO. Moreover, the possibility to nullify the effects of the negative externality through rejection further increases their willingness to reject. We show that the latter result is better explained by a model of consequentialist moral concerns toward the NGO rather than deontological morality about own actions.

Keywords: three-player ultimatum game, moral reasoning, experiment.

JEL code: C72, C91, D6

Acknowledgements: We thank the ESRC (NIBS Grant ES/K002201/1) for financial support. We also thank Werner Güth, Francesco Guala, Marta Barazzetta, Jana Willrodt, Felix Kölle, Matteo Ploner and attendants at the NIBS internal workshop at the University of Warwick and at SEET 2017 Workshop for useful comments and suggestions.

1 University of Milan, Department of Economics, Management, and Quantitative Methods (DEMM), via Conservatorio 7, 20122 Milan, Italy. E-mail: sandro.casal@unimi.it

2 University of East Anglia, School of Economics, Norwich Research Park, Norwich, Norfolk, NR4 7TJ, UK. E-mail: f.fallucchi@uea.ac.uk

3 Corresponding author. University of Bonn, Institute for Applied Microeconomics (IAME), Adenauerallee 24 – 42, 53113, Bonn, Germany. E-mail: simone.quercia@uni-bonn.de

1. Introduction

Under what conditions are people willing to sacrifice possible gains from trade to punish immoral behavior toward a third party? The investigation of this question has been at the heart of a flourishing and lively debate in economics. Falk and Szech (2013) provided evidence that market interactions tend to lower moral values compared to non-market transactions. In contrast, Bartling et al. (2015) find that socially responsible market behavior mitigates the potential damages of negative externalities: when facing multiple product offers, consumers are more inclined to buy 'ethical' products even if this entails paying higher prices. More recently, Kirchler et al. (2016) investigated which type of intervention reduces the extent of immoral behavior toward a third party in market and non-market transactions.

An open question, which we investigate in this paper, is whether subjects who refuse to engage in immoral transactions are motivated by the willingness to punish the perpetrators of immoral actions or by the intention of reducing or eliminating the negative externality on the third party. In the market experiments cited above, due to repeated interactions, both motives may be present. However, this distinction is important for understanding individuals' motivation in fighting unethical behavior. For example, product boycotts are common practices of attempts to fight companies' socially irresponsible or immoral behavior. Boycotters bear the cost of avoiding the product consumption and often, the cost of finding an alternative, more 'ethical' product. In these cases, their behavior may be driven by two motives: the willingness not to engage in a transaction with the company just because its behavior is perceived morally wrong or the attempt to restore justice toward the victims of the immoral conduct by pushing the company to produce in a more ethical or sustainable way. While the former may be related to subjects' deontological morality, as subjects may be willing to pursue what is morally right no matter the consequences of their action, the intention to restore justice seems motivated by a consequentialist view of own actions and may also play a key role on subjects' decision to boycott (Sandel, 2010; White et al., 2012).

We aim to distinguish these two motivations by analyzing subjects' behavior in a modified version of the three-person ultimatum bargaining game proposed by Güth and Van Damme (1998) (henceforth GVD). In particular, we are interested in how responders perceive and react to proposers' morality (or lack thereof). We decided to build our experiment on an ultimatum game because, beside the simplicity of the game structure, it has been frequently used to study fairness considerations and moral concerns (see Güth and Kocher, 2014, for a review).

In our experiment, the proposer can increase their initial endowment by *taking* an amount from a dummy player (see List, 2007 and Bardsley, 2008). Then, the proposer offers a fraction of their endowment to the responder choosing among two possible offers, high or low, which allows us to clearly distinguish between kind and unkind offers (see Bolton and Zwick, 1995 and Falk et al., 2003). Finally, the responder decides whether to accept or reject the offer.

The payoff consequences of a rejection vary across treatments. The first and second treatments mimic two of the information conditions of GVD, i.e., responders decide to accept or reject an offer having either no information or full information on the share of the other two players. In these two treatments, a rejection has payoff consequences only for the proposer and the responder. In case of rejection both proposer and responder get zero, while in case of acceptance the proposed offer is implemented. In our third treatment, a rejection has instead payoff consequences also for the dummy player. This treatment mirrors the full information treatment except that rejecting an offer restores the initial allocation to the dummy player while still punishing the proposers by setting their payoff to zero.

In the no information treatment, responders are not aware of the amount taken from the dummy player and hence on the amount the proposer is keeping for themselves. This treatment serves as a benchmark to assess responders' reactions to 'revealed' morality of proposers in the full information treatment. Without information, responders may be inclined to accept offers more frequently hiding behind their ignorance on the faith of the dummy player. This effect would be consistent with a self-serving notion of fairness and morality (see Dana et al., 2007).

The difference between the second and third treatment stands only in the payoff consequences for the dummy player. Hence, the comparison between our second and third treatment investigates our main research question and reveals the moral motivation for a rejection, that is, whether rejections are only motivated by the desire to punish proposers or whether the consequences for the dummy player also play an important role. To guide our analysis of moral reasoning across these two treatments, in Section 2.3, we formulate and contrast predictions from three different models that may explain behavior in our setup: a standard model of social preferences, a consequentialist model of moral concerns, and a model of deontological morality.

Our dummy player is an NGO (Médicins Sans Frontières - Doctors without borders) and its initial endowment is a donation provided by the experimenter to the NGO (see Kirchler et al., 2016, for a similar approach). We consider the act of taking from the NGO as an indication of the degree of morality of proposers for two main reasons: first, acts of taking are

generally perceived as less socially appropriate than corresponding acts of giving (Krupka and Weber, 2013); second and more important, since we chose an NGO that conducts medical humanitarian activities in emergency areas, any amount not donated reduces the fundraise for Doctors without Borders, which, in turn, makes the accomplishment of its goals more difficult (primarily sustaining people in serious medical conditions in emergency areas).

Our findings reveal that uninformed responders behave as if they had no moral concerns for the NGO. However, they take into account the negative externality on the NGO in the treatments where this information is revealed: the larger the amount taken from the NGO, the more likely the offer is rejected. Finally, in our last treatment we find evidence that the possibility to restore the initial donation makes responders more willing to reject a kind offer compared to situations where justice restoration is not possible. This is not true for unkind offers as those are frequently rejected both with and without the possibility of justice restoration. We speculate that the opportunity of nullifying the payoff consequences of an injustice makes responders feel more responsible for the NGO as they are more willing to forgo their share in order to restore the initial donation. Going back to the example of product boycott, consumers may be more willing to boycott firms who produce unethically if they can see consequential beneficial results of their actions on the third parties who are bearing the unethical consequences of the firm behavior.

The remainder of the paper is structured as follows. Section 2 presents the experimental design, behavioral predictions, and experimental procedures. Section 3 contains the data analyses and results. Section 4 concludes and discusses implications of our findings.

2. Experimental design, behavioral predictions and procedures

2.1 Experimental design

In our design, the proposer (P) is initially endowed with €9 (E_P), the responder (R) with €0 (E_R) and an initial donation of €5 is allocated to the NGO (E_{NGO}). The experiment consists of three phases.

PHASE 1 - P takes a discrete sum t from the initial donation to the NGO, where €0 ≤ t ≤ €5.

PHASE 2 - P makes an offer o, of either €5 or €1, to R.

PHASE 3 - R chooses to accept or reject the offer. In case of acceptance the amounts are distributed as proposed; in case of rejection both P and R receive €0.

The three phases are common knowledge. We employ three treatments to investigate our research questions. In the first treatment, NO-INFO, the amount taken by P from the NGO is concealed. Hence, R makes a decision of accepting or rejecting an offer without knowing the outcome of PHASE 1. In our second treatment, INFO, responders are informed about the amount taken from the NGO before making a decision on the acceptance of the offer. Our third treatment, INFO-RESTORE, is like the INFO treatment except that the rejection of an offer implies different consequences for the NGO. While a rejection in INFO (and NO-INFO) sets the final payoff of P and R to zero and does not affect the NGO's payoff, in INFO-RESTORE a rejection restores the initial donation to the NGO to €5. Table 1 shows the consequences of responders' actions for our three treatments.[1]

Table 1. Payoffs

	NO-INFO and INFO		**INFO-RESTORE**	
	R accepts	R rejects	R accepts	R rejects
P	$E_P + t - o$	0	$E_P + t - o$	0
R	$E_R + o$	0	$E_R + o$	0
NGO	$E_{NGO} - t$	$E_{NGO} - t$	$E_{NGO} - t$	E_{NGO}

The subgame-perfect Nash equilibrium of the game under standard preferences is the same in all treatments: R accepts all offers while P takes the maximum allowed amount from the NGO and offers the smallest possible amount to R. However, if subjects have social preferences or moral concerns they might behave differently across the three treatments. In the next section, we analyze three theoretical models reflecting different types of social or moral concerns that subjects may have in our setting and derive predictions on responders' behavior for each model.

2.2 Behavioral predictions

In this section, we derive predictions for responders' behavior in INFO and INFO-RESTORE. Predictions for NO-INFO are analogous to INFO except that responses rely on beliefs about proposers' taking rather than observation of the amount taken. We focus on responders as their behavior is the main interest of our research question and they are not

[1] A treatment similar to our INFO, where the dummy player may receive a positive amount in case of Responder's rejection, has been tested by Pfister and Böhm (2012).

affected by strategic considerations. We derive and compare behavioral predictions from three different models that may explain behavior in our setup.

The first model we consider is inequity aversion as introduced by Fehr and Schmidt (1999). In our setting, this model assumes that responders suffer costs from advantageous and disadvantageous inequity both toward the proposer and toward the NGO. We use this model as our starting point because it is a canonical model of social preferences that has been used previously to explain behavior in ultimatum games (see Fehr and Schmidt, 1999 and Fehr and Schmidt, 2003). However, it has been also noted that this model does not capture some evidence from three-player ultimatum games as responders do not seem to display concerns toward dummy players (see GVD and Guth et al., 2007). Moreover, previous three-player ultimatum games often used other experimental participants as dummy players instead of an NGO. While for these studies it is sensible to assume that subjects care about inequity toward all other agents since they are similar at the beginning of the experiment, it might not be very plausible in our setup where the NGO has different needs and characteristics compared to active players. Therefore, in our second model responders have inequity concerns toward the proposer, but have consequentialist moral concerns toward the NGO. In particular, they suffer a cost which is increasing in the distance between the NGO's payoff and the maximal donation. Finally, we introduce a model of deontological or Kantian morality[2] which assumes that subjects follow the moral imperative: "Do not accept 'immoral' offers". In particular, they have inequity concerns toward the proposer as before, but bear a moral cost if they accept an offer from a proposer who has taken a positive amount from the NGO. This cost is independent of any consequences that the offer and a potential rejection may have and only depends on own actions.

In what follows, we report behavioral predictions for differences across INFO and INFO-RESTORE and explain the intuition behind each prediction. We report formal proofs and a more detailed discussion of the models in Appendix C.

2.2.1 Inequity aversion (Fehr and Schmidt, 1999)

Following Fehr and Schmidt (1999), we define the responder's utility function as:

[2] Throughout the paper, we use interchangeably the adjectives "deontological" and "Kantian". In both cases, we mean that people suffer a moral cost if own actions deviate from their moral duties or imperatives.

$$U_R = \pi_R - \frac{\alpha}{2}[\max(\pi_P - \pi_R\,, 0) + \max(\pi_{NGO} - \pi_R\,, 0)] - \frac{\beta}{2}[max(\pi_R - \pi_P\,, 0) + max(\pi_R - \pi_{NGO}\,, 0)]$$

where $0 \leq \beta < 1$ and $\alpha \geq \beta$ are the inequality parameters and π_P, π_R, π_{NGO} are the payoffs of the three players.

According to this utility function we derive the following propositions:

PROPOSITION 1a. *For offers of €5 acceptance rate is 100% both in INFO and INFO-RESTORE when* $t \leq 3$. *Acceptance is more likely in INFO-RESTORE than in INFO when* $t > 3$.

Proofs. see Appendix C.1.a.

For low levels of t, the benefit of establishing equality (from the responder point of view) between the responder and the proposer via a rejection is offset by the sum of the monetary cost of a rejection plus the disadvantageous inequality toward the NGO that a rejection would generate. Hence, for $t \leq 3$ offers of €5 will be always accepted.

For high values of t, the intuition behind the predicted treatment difference between INFO and INFO-RESTORE is that in case of acceptance in INFO the disadvantageous inequity toward the proposer and the advantageous inequity toward the NGO are so high that for some values of α and β the responder is willing to reject to have equality with the proposer and small (or no) disadvantageous inequality toward the NGO. In contrast, in INFO-RESTORE, given that a rejection restores the initial donation to the NGO the disadvantageous inequity toward the NGO is maximal regardless of t. Therefore, the cost of rejection is high enough to induce the responder to accept offers more frequently than in INFO.

PROPOSITION 1b. *For offers of €1 acceptance happens only if subjects care very little about disadvantageous inequality. Given all possible values of α (concern for disadvantageous inequality) acceptance is more likely in INFO-RESTORE than in INFO.*

Proofs. see Appendix C.1.b.

The intuition behind the first part of the proposition is trivial: as the inequality generated by the acceptance of an offer of €1 is very high, the responder needs to care very little about disadvantageous inequity to accept such an offer. Hence, acceptance will happen only for very low levels of α.

The intuition for the second part is similar to the case of €5 offers as a rejection in INFO-RESTORE implies more inequity than in INFO because of the restoration option. However,

cases where the model predicts rejection in INFO and acceptance in INFO-RESTORE happen for small intervals of α. For most plausible values of α as estimated empirically by previous literature (see, e.g., Bellemare et al., 2008 and Blanco et al., 2011), offers of €1 are rejected in both treatments. For this reason, we anticipate higher likelihood to detect treatment differences in our data for offers of €5 rather than for offers of €1.

To summarize, according to inequity aversion (Fehr and Schmidt, 1999), acceptance rates in INFO are predicted to be lower than in INFO-RESTORE. This pattern should be stronger for offers of €5 than for offers of €1.

2.2.2 Inequity aversion toward the proposer and consequentialist morality toward the NGO

The second model we propose and analyze is a consequentialist model where responders care about inequity toward the proposer, but also cares about the distance between the NGO's payoff and the maximal possible donation. To characterize this model, we specify the following utility function:

$$U_R(\pi_P, \pi_R, \pi_{NGO}) = \pi_R - \alpha[\max(\pi_P - \pi_R\,, 0)] - \beta[\max(\pi_R - \pi_P\,, 0)] - \gamma(5 - \pi_{NGO})$$

where $0 \leq \beta < 1$, $\alpha \geq \beta$, and $0 < \gamma < 1$. The assumptions on γ guarantee that the responder is not spiteful toward the NGO ($\gamma > 0$) and that they care about their payoff relatively more than about the NGO's payoff ($\gamma < 1$). Given the utility function above we specify the following propositions:

PROPOSITION 2a. *For offers of €5 acceptance rate is 100% both in INFO and INFO-RESTORE when* $t \leq 1$. *Acceptance is more likely in INFO than in INFO-RESTORE when* $t > 1$.

Proofs. see Appendix C.2.a.

The intuition behind the first part of the proposition is that, for low levels of taking, the moral concern for the NGO and inequity costs are not sufficiently high to lead to a rejection of a kind offer.

With respect to the second part, it is straightforward to see that a rejection in the INFO-RESTORE eliminates the moral cost toward the NGO that acceptance would generate, while in INFO this cost is present in case of acceptance as well as in case of rejection, making a rejection less appealing.

PROPOSITION 2b. *For offers of €1 acceptance happens only if subjects care very little about disadvantageous inequality. Given all possible values of α (concern for disadvantageous inequality) acceptance is more likely in INFO than in INFO-RESTORE.*

Proofs. see Appendix C.2.b.

The intuition for the first part of the proposition is similar to the standard Fehr and Schmidt (1999) model described above and implies that we should expect most of the €1 offers to be rejected if subjects care about disadvantageous inequity to an extent similar to the one reported in previous studies (see, e.g., Bellemare et al., 2008 and Blanco et al., 2011).

As in the case of offers of €5, the second part of the proposition derives from the fact that the moral cost toward the NGO is always present for positive taking in INFO while it disappears in INFO-RESTORE if the responder rejects the offer.

To sum up, according to this second model, acceptance rates in INFO are predicted to be higher than in INFO-RESTORE. As before, this effect should be more detectable for offers of €5 than for offers of €1. Notice that this model offers the opposite prediction compared a standard inequity aversion model regarding the direction of the treatment difference between INFO and INFO-RESTORE.

2.2.3 Inequity aversion toward the proposer and Kantian morality toward the NGO

Finally, we present a simple model where we assume that the responder does not have moral concerns about the NGO payoffs but rather about own actions. In particular, we assume that the action of accepting an offer generated after the proposer has taken positive amounts from the NGO is seen as 'immoral' and hence it creates a psychological cost no matter the extent of taking and the consequences for the NGO. This type of morality that we label 'Kantian' is unconditional, in the sense that the moral imperative of not committing an immoral action (accepting an immoral deal) does not depend on the consequences of the action.[3] With this in mind we specify the following utility function:

$$U_R(\pi_P, \pi_R, K) = \pi_R - \alpha[\max(\pi_P - \pi_R\,, 0)] - \beta[\max(\pi_R - \pi_P\,, 0)] - \mathbb{1}_{t>0}\mathbb{1}_{Accept}K$$

[3] To the best of our knowledge, the only model in the economic literature that incorporates Kantian morality into a utility function is Alger and Weibull (2013). Their model is based on symmetric games and is more general than our formulation (see Miettinen et al. (2017) for an experimental test of the model). Our aim is not to provide a general theory of Kantian morality, but to assume a Kantian moral imperative, i.e., "do not accept an 'immoral' offer", that is reasonable in our setup and describes the logic of our initial research question. In this sense, our model is similar in spirit to the procedural fairness model in Dold and Khadjavi (2017).

where $0 \leq \beta < 1$, $\alpha \geq \beta$, $K > 0$, $\mathbb{1}_{t>0}$ is an indicator function that takes value 1 when the proposer has taken a positive amount from the NGO and 0 otherwise, and $\mathbb{1}_{Accept}$ takes value 1 if the responder accepts and 0 otherwise. Given the specification above we derive the following proposition:

PROPOSITION 3. *Acceptance is equally likely in INFO-RESTORE and in INFO for both offers of €1 and €5.*

Proof. see Appendix C.3.

The proposition follows directly from our specification of the utility function. It stems from the fact that the mere act of accepting an 'immoral' offer constitutes a cost, no matter the final payoff for the NGO. Given that the final payoff for the NGO is only difference across INFO-RESTORE and INFO, the model predicts no differences across the two treatments.

2.3 Procedures

The Experiment was conducted at LabSi (University of Siena, Italy) using z-Tree (Fischbacher, 2007). A total of 174 subjects were recruited for 11 sessions (3 for the NO-INFO treatment, 4 for the INFO and INFO-RESTORE treatments). Upon their arrival to the lab subjects were seated in cubicles that separated them visually to protect anonymity. Instructions, that we report in Appendix A (translated from Italian), were read aloud by the experimenter. After reading the instructions, the experimenter highlighted that the amount left to the NGO would be donated a few weeks after the experiment and that subjects would receive an email with the receipt of the total donation. Average payment for subjects was €6.85, including a €3 show-up fee, for an average session duration of 30 minutes. As NGO, we chose "MSF - Doctors without Borders" because it is a well-known and esteemed organization in Italy where the experiment was conducted and it does not have any strong religious or political affiliation. Moreover, as mentioned above, the organization has a clear humanitarian mission and this gives to the act of taking from the NGO the moral connotation needed for our research question. The total donation to "MSF - Doctors without borders" was €220 (out of a potential maximum amount of €435).

In our experiment, in PHASE 1 and 2 proposers played the taking game and chose the amount to offer to responders, respectively. In PHASE 3, we adopted a procedure that uses a

variant of the Random Lottery incentive system (see, e.g., Cubitt et al., 1998) to elicit responders' choices for different paths of play. In particular, all responders received the information about the offer made by each proposer in a sequence of screens, and made the choices knowing that only one would have been relevant for the final payoff. After PHASE 3 the computer randomly paired all subjects in the session and calculated the outcome of the game.

This procedure shares some similarities with the strategy method (Selten, 1967) and with the conditional information lottery (CIL, Bardsley, 2000) as it allows us to observe choices of responders at different nodes of the game. We think, however, that this approach has some advantages in our setup compared to the mentioned elicitation methods. One common critique of the strategy method is that it elicits choices for all information sets and hence also for nodes that may be unlikely to be reached. This implies that subjects may give different weights to their decisions. This problem can become serious if subjects attach zero probability to a certain node as the corresponding decision would not be incentive-compatible. In the CIL, this concern is slightly less problematic as the researcher can select fictitious nodes that are likely to occur in actual play. However, this implies that the researcher needs to know not only which nodes of the game are likely to occur but also which nodes the subjects believe to be likely to occur. This proves to be difficult to achieve in many cases. Our method solves these concerns as responders make decisions in nodes of the games that have been *actually* reached by the play of proposers and hence each decision that responders take has *positive* and *equal* probability of being implemented.

It is also worth noting that the number of choices taken by each responder in our setting would have been different across treatments had we used a standard implementation of the strategy method. This is because in NO-INFO the information sets are less than in the other two treatments. Our method keeps the number of choices per responder similar across treatments by asking responders to make a choice for each proposers' action.[4]

There are also some disadvantages of our method: first of all, each session constitutes an independent observation as each session produces a different set and sequence of proposers' actions. Hence, we are forced in the data analysis to use regression models with clusters at individual and session level when we analyze responders' behavior. This is not needed in the

[4] In our experiment, all sessions but one were composed by 16 subjects, which implies 8 choices per responder. Only one session in INFO-RESTORE was composed by 14 subjects which involves 7 choices per responder.

strategy method where each individual's strategy constitutes an independent observation. Second, to observe choices of responders at different nodes, we need a small enough strategy space to generate the necessary variation. Our game, structured as a mini-ultimatum game, achieves this generating enough 'cases' to create within-subject counterfactual comparisons in responders' behavior.[5]

3. Results

We divide our results section in two sub-sections. In Section 3.1 we investigate our main research question analyzing behavior of responders across our three treatments. In Section 3.2 we report results on proposers' behavior. In analyzing responders' behavior, given that they face a sequence of offers that is session-dependent we implement only regression analysis (mixed effects models) with clusters at individual and session level. With regard to proposers' behavior, we treat each offer and taking decision as independent.

3.1 Responders' behavior

3.1.1 The effect of kind or unkind offers on acceptance

Result 1. *In all treatments, responders are significantly less likely to accept unkind offers than kind offers. Acceptance is highest in the NO-INFO treatment.*

Responders in NO-INFO almost always accept a €5 offer (95%) while they accept a €1 offer 50% of the times. The acceptance rate drops to 88% and 84% for €5 offers and to 28% and 26% for €1 offers in INFO and INFO-RESTORE, respectively. In Table 2, we report the estimates of mixed effect probit models investigating differences in acceptance depending on the offer received. In each model we regress an acceptance dummy (equal to one if the offer is accepted) on a €5 dummy (equal to one if the offer received is €5) for each treatment separately. The coefficients of the €5 dummy are positive and highly significant in all three models indicating that responders accept more often a "kind offer" of €5 than an "unkind offer" of €1 in all treatments.

[5] For a more general discussion and survey on the differences between the strategy method and the direct-response method see Brandts and Charness (2011). Güth et al. (2001) provide an example where the direct-response and the strategy method elicitations lead to different behavior in mini ultimatum games.

Table 2. The relation between acceptance and offer received

	NO-INFO (1)	INFO (2)	INFO-RESTORE (3)
€5 offer	2.01***	2.69***	2.03***
	(0.31)	(0.33)	(0.27)
Constant	-0.01	-0.91***	-0.80***
	(0.22)	(0.29)	(0.27)
N	192	256	241

Notes: mixed-effects probit with random intercepts at the individual and session level. Dependent variable = 1 if the offer is accepted and = 0 otherwise. * significant at 10%, ** significant at 5%, *** significant at 1%.

It is interesting to note that responders are significantly more optimistic in NO-INFO compared to the two full information treatments regarding the amount taken from the NGO. Non-incentivized beliefs about the modal amount taken (elicited before responders saw the actual amounts taken) are on average €2.50 in NO-INFO, €3.44 in INFO and €2.84 in INFO-RESTORE; a comparison of beliefs of the NO-INFO treatment versus the two full information treatments shows a statistically significant difference (MWU-test, $p = 0.037$). As subjects had no incentive to report 'correct' beliefs, this suggests that responders may have reported deflated beliefs in NO-INFO to justify their high acceptance rate.

So far, we have neglected the role of the information on the amount taken which likely plays an important role on the decision to accept. Hence, in the next subsections we analyze treatment differences taking into account the information on the amount taken.

3.1.2 The effect of information on acceptance

In this section, we analyze the effect of 'revealed' information on responders' behavior. For this purpose, we concentrate on the comparison between the NO-INFO and the INFO treatments since the only difference across these two treatments is the concealment of information in NO-INFO.

Result 2. *When responders are fully informed, their willingness to punish increases with the amount taken from the NGO.*

In Table 3 we estimate the effects of disclosing the information about the amount taken on the acceptance rate using linear mixed effect models.[6] Our dependent variable is a dummy that takes value one if the offer is accepted and zero otherwise. Our independent variables are the amount taken, a dummy for the INFO treatment and the interaction term between the amount taken and the treatment dummy. We estimate the model both pooling all offers (Model (1)) and separating between kind and unkind offers (Models (2) and (3)). In the NO-INFO treatment, the amount taken from the NGO is concealed and hence cannot influence responders' decisions. All three regressions confirm this, as the coefficient of Amount taken is not significantly different from zero. Disclosing the information about the amount taken from the NGO has a significant effect on acceptance. When considering all offers (Model (1)), responders tend to accept more compared to NO-INFO if proposers have not taken any positive amount from the NGO (coefficient of INFO). However, the likelihood of accepting decreases significantly with the increase in the amount taken (see interaction term). In Models (2) and (3) we analyze kind and unkind offers separately. Interestingly, when proposers do not take any positive amount from the NGO, the acceptance rate of offers of €1 is not significantly different between the two treatments, while it is for offers of €5 (see coefficients of INFO). However, for both unkind and kind offers the higher the amount taken, the lower the willingness to accept (see interaction terms). This decreasing pattern of acceptance in INFO was predicted by all models presented in Section 2.2.

[6] Although our dependent variable is dichotomous, in these specifications we use linear models rather than non-linear probit or logit models because we estimate regressions with interaction terms. The reason is that, as noted by Ai and Norton (2003) the sign of the coefficient of an interaction term estimated through non-linear models "*does not necessarily indicate the sign of the interaction effect*" since "*the interaction effect may have different signs for different values of covariates*" (page 124).

Table 3. The effect of information on acceptance.

	All offers	€1 offers	€5 offers
	(1)	(2)	(3)
Amount taken	0.02	-0.01	-0.01
	(0.02)	(0.02)	(0.01)
INFO	0.21**	-0.03	0.17**
	(0.09)	(0.13)	(0.09)
INFO × Amount taken	-0.11***	-0.07***	-0.08***
	(0.02)	(0.03)	(0.02)
Constant	0.69***	0.52***	0.98***
	(0.06)	(0.09)	(0.07)
N	448	184	264

Notes: mixed-effects linear model with random intercepts at the individual and session level. Dependent variable takes value 1 if the offer is accepted and zero otherwise.
* significant at 10%, ** significant at 5%, *** significant at 1%.

3.1.3 The effect of justice restoration on acceptance

To analyze the effects of the opportunity of justice restoration we focus on the comparison between INFO and INFO-RESTORE, where the only difference across treatments is the consequence of a rejection for the NGO.

Result 3. *For positive amounts taken from the NGO, responders reject a €5 offer more often when the rejection restores the initial donation.*

In both full information treatments, roughly 60% of the subjects always reject €1 offers. However, if we consider €5 offers, around 80% of responders in INFO are always willing to accept them compared to 55% in INFO-RESTORE. This suggests that responders take into consideration the opportunity of restoring the consequences of an immoral action of the proposers, especially when they receive an offer of €5. In Figure 1, we compare how responders' behavior changes between INFO and INFO-RESTORE for each amount taken from the NGO. In the left panel, showing the acceptance rate for €1, we observe the highest acceptance rate for both treatments when the proposer does not take any amount from the NGO.

Given the low number of €1 offers and the high rate of rejections it is difficult to find other visual patterns.

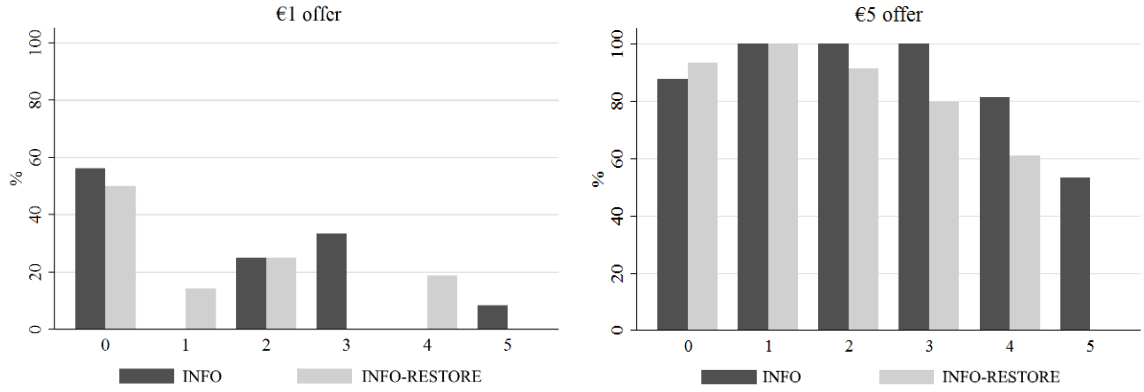


Figure 1. Acceptance rate based on amount taken in PHASE 1.

The more interesting comparison is for offers of €5 in the right panel of Figure 1. The acceptance rate of a €5 offer is high in INFO, close to 100%, if proposers have taken €3 or less; it decreases to 80% when the proposers have taken €4 and to around 50% in case of €5 taken. Thus, subjects are willing to forgo their payoff to punish proposers, although this appears evident only when proposers take high amounts from the NGO. In INFO-RESTORE the drop starts at €2, and decreases faster compared to INFO, indicating that the possibility to restore the ex-ante situation has a sizeable effect on subjects' willingness to punish.

In Table 4, we report results of regression analyses that compare the two treatments. We restrict our estimates to the cases where the amount taken from the NGO is greater than zero as the restore option does not have an effect when the amount taken is equal to zero. We regress an acceptance dummy on the treatment dummy INFO-RESTORE and on the amount taken from the NGO. Model (1) is estimated for all offers, while Model (2) for offers of €5.[7] In Model (1), we find that the amount taken significantly decreases the acceptance rate, but the acceptance rate is not different between the INFO and INFO-RESTORE. When we analyze €5 offers, we still find that the amount taken has a significant effect on the likelihood of acceptance but on top of that responders tend to accept significantly less when the restore option is

[7] We also estimated the same specification with the subset of offers of €1, but due to the low number of offers and the clusters at session and individual level, the log-likelihood estimation does not converge. Hence, we run two additional sets of estimations that we report in Appendix B. The first set uses a mixed effect probit model with clusters only at session level (Table B1). The second set of estimates uses linear mixed effect models instead of mixed effects probit models and maintains clusters both at the individual and at session level (Table B2). See discussion of the results of these additional estimations in the main text.

available (Model (2)). Other specifications of this model, reported in Appendix B, which allow us to estimate also the effect on offers of €1, reveal that the restore option lowers significantly the acceptance rate for offers of €5 but not for offers of €1. Taken together, these results give stronger support to our second model reported in Section 2.2, that is, a model where subjects have inequity concern toward the proposer and consequentialist morality toward the NGO.

Table 4. The effect of the restore option.

	All offers (1)	€5 offers (2)
INFO-RESTORE	-0.12	-3.52**
	(0.29)	(1.40)
Amount taken	-0.47***	-1.85***
	(0.07)	(0.42)
Constant	1.93***	10.87***
	(0.30)	(2.57)
N	442	316

Notes: mixed-effects probit with random intercepts at the individual and session level. Dependent variable takes value 1 if the offer is accepted and zero otherwise. * significant at 10%, ** significant at 5%, *** significant at 1%.

3.2 Proposers' behavior

Result 4. *The distribution of amounts taken by proposers is different between the NO-INFO and the two full information treatments.*

Figure 2 shows the distribution of amounts taken from the NGO across treatments. The average amount taken in NO-INFO is €3.04, with €5 being the modal choice, although more than 20% of proposers do not take any positive amount from the NGO. In the two full information treatments, the amount taken is lower averaging €2.78 in INFO and €2.29 in INFO-RESTORE with modal choices of €2 and €3, respectively. A $\chi 2$-test rejects the null hypothesis of no differences between the three distributions ($p = 0.039$). Pairwise comparisons show no differences between the two full information treatments ($\chi 2$-test, $p = 0.191$) while both of them are significantly different compared to NO-INFO (INFO vs. NO-INFO $\chi 2$-test, $p = 0.081$; INFO-RESTORE vs. NO-INFO $\chi 2$-test, $p = 0.029$). This suggests that proposers modify significantly their behavior in response to the different information structure of NO-INFO and

INFO strategically anticipating responders' reactions. They do so also between INFO and INFO-RESTORE as suggested by the lower average amount taken in INFO-RESTORE but to a smaller extent and not significantly so.

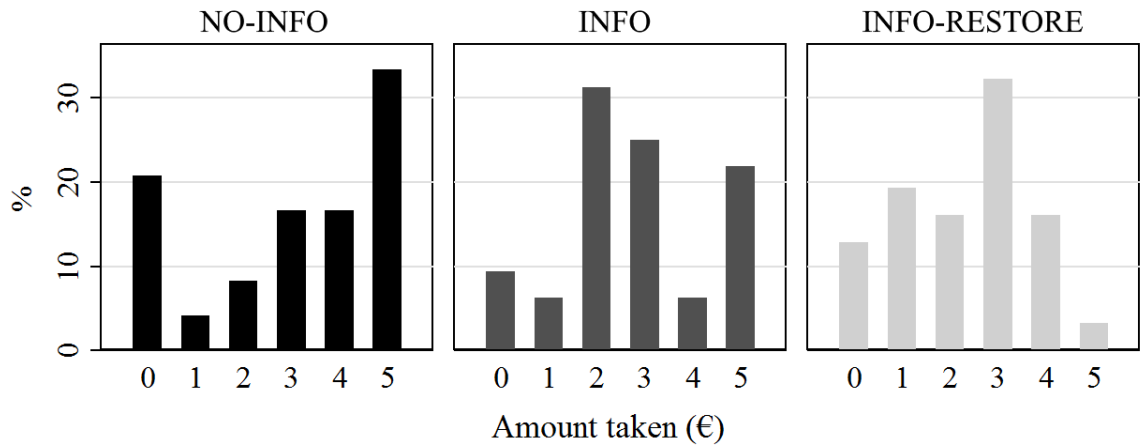


Figure 2. Distribution of amount taken (€) across treatments

Result 5. *The average amount offered is not significantly different across the three treatments.*

On average 64% of the proposers across treatments offer €5, while 36% offer €1. Figure 3 shows the share of €1 offers across treatments. This share is highest in NO-INFO and lowest in INFO-RESTORE. However, pairwise statistical comparisons do not reveal significant differences across treatments (χ^2-tests, all $p > 0.157$).

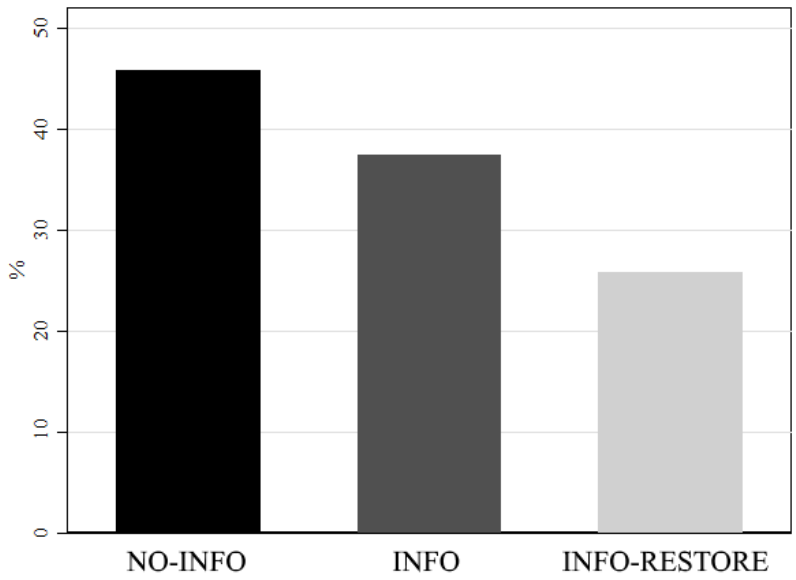


Figure 3. Shares of 1€ offers across treatments

In line with previous three-person ultimatum game experiments (Güth et al., 2007) a small percentage of proposers (< 5%) offered a larger share to responders than to themselves. Proposers in NO-INFO offer on average 26% of their resources (initial endowment plus the amount taken from the NGO). This share increases to 30% in INFO and 35% in INFO-RESTORE.[8] Binary comparisons reveal a significant difference only between NO-INFO and INFO-RESTORE (χ^2-tests, $p = 0.049$; all other comparisons, $p > 0.136$).

4. Conclusion

In this paper, we have studied the role of moral concerns in three-player ultimatum bargaining. Our results show that moral concerns do play an important role in the decisions of both parties (proposers and responders). Proposers tend to take more from the NGO when this action is not revealed to responders. This suggests that proposers anticipate responders' reactions and take strategically less from the NGO under full information to avoid rejections. Hence, policy proposals to promote transparency of firms' moral conduct toward consumers may be beneficial to reduce potential harm on third parties. In relation to this, it is not surprising that companies or individuals, whose immoral conducts have been made public, try subsequently to implement and promote moral and fair initiatives (Bowie and Dunfee, 2002).

The importance of moral concerns is even clearer when analyzing responders' behaviors: they are willing to punish, against their own personal interests, an immoral act toward the NGO, and they do so even more when the punishment restores justice. Our results are best predicted by a model that assumes inequity aversion toward the proposer and consequentialist morality toward the NGO. Predictions from standard inequity aversion (Fehr and Schmidt, 1999) or a mixture of inequity aversion and deontological morality (see Section 2.2) are not borne out by data. One important question for future research is to investigate the role played by the sense of responsibility that responders feel for the consequences of their action for the NGO. In our experiment, when a responder in the INFO-RESTORE decides to reject the offer, they are fully responsible for the NGO and certain that the situation of the NGO would improve due to their action. In reality, justice restoration may not be such an easy process. For this reason, one interesting direction might be to look at intermediate situations where the remedy to the immoral actions depends on reaching a critical mass of punishment among responders.

[8] Given the structure of our UG game, the minimum share is 7% when the proposer takes €5 from the NGO and offers €1 to the responder. The maximum share is 56% when the proposer takes €0 and offers €5.

References

Ai, C., Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics letters*, *80*(1), 123-129.

Alger, I., Weibull, J.W. (2013). Homo moralis - preference evolution under incomplete information and assortative matching. *Econometrica*, 81(6), 2269-2302.

Bardsley, N. (2000). Control without deception: Individual behaviour in free-riding experiments revisited. *Experimental Economics*, 3(3), 215-240.

Bardsley, N. (2008) Dictator game giving: altruism or artifact? *Experimental Economics*, 11 (2), 122-133.

Bartling, B., Weber, R., Yao, L. (2015). Do markets erode social responsibility? *Quarterly Journal of Economics*, 130(1), 219-266.

Bellemare, C., Kröger, S., Van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica*, 76(4), 815-839.

Blanco, M., D. Engelmann, and H. Normann (2011): A within-subject analysis of other-regarding Preferences. *Games and Economic Behavior*, 72, 321-338.

Bolton, G.E., Zwick, R. (1995). Anonymity versus punishment in ultimatum bargaining. *Games and Economic behavior*, 10(1), 95-121.

Bowie, N.E., Dunfee, T.W. (2002). Confronting morality in markets. *Journal of Business Ethics*, 38(4), 381-393.

Brandts, J., Charness, G. (2011). The strategy vs. the direct-response method: a first survey of experimental comparisons. *Experimental Economics*, 14(3), 375-398.

Cubitt, R. P., Starmer, C., Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, 1(2), 115-131.

Dana, J., Weber, R. A., Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1), 67-80.

Dold, M., and Khadjavi, M. (2017). Jumping the queue: An experiment on procedural preferences. *Games and Economic Behavior*, 102, 127-137.

Falk, A., Fehr, E., Fischbacher, U. (2003). On the nature of fair behavior. *Economic Inquiry*, 41(1), 20-26.

Falk, A., Szech, N. (2013). Morals and markets. *Science*, 340, 707–711.

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

Fehr, E., Schmidt K.M. (2003) Theories of fairness and reciprocity: evidence and economic applications. In: Dewatripont M., Hansen L.P., Turnovsky S.J. (eds). *Advances in economics and econometrics: theory and applications*, Eight World Congress, Volume I. Cambridge University Press, Cambridge, pp 208–257.

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

Güth, W., Huck, S., Müller, W. (2001). The relevance of equal splits in ultimatum games. *Games and Economic Behavior*, *37*(1), 161-169.

Güth, W., Kocher, M. (2014). More than thirty years of ultimatum bargaining experiments: motives, variations, and a survey of the recent literature. *Journal of Economic Behavior and Organization*, 108(C), 396-409.

Güth, W., Schmidt, C., Sutter, M. (2007). Bargaining outside the lab. A newspaper experiment of a three-person ultimatum game. *The Economic Journal*, 117(518), 449-469.

Güth, W., Van Damme, E. (1998). Information, strategic behavior, and fairness in ultimatum bargaining: An experimental study. *Journal of Mathematical Psychology*, 42(23), 227-247.

Kirchler, M., Huber, J., Stefan, M., Sutter, M. (2016). Market design and moral behavior. *Management Science*, 62(9), 2615-2625.

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

List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, 115(3), 482-493.

Miettinen, T., Kosfeld, M., Fehr, E., Weibull, J.W. (2017). Revealed preferences in a sequential prisoners' dilemma: a horse-race between five utility functions. CESifo Working Paper Series No. 6358.

Pfister, H.R., Böhm, G. (2012). Responder feelings in a three-player three-option ultimatum game: Affective determinants of rejection behavior. *Games*, 3(1), 1-29.

Sandel, M. J. (2010). Justice: What's the right thing to do? Macmillan.

Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments, in H. Sauermann (ed.), *Beiträge zur experimentellen Wirtschafts-Forschung*. Tübingen: JCB Mohr, 136–168.

White, K., MacDonnell, R., Ellard, J.H. (2012). Belief in a just world: Consumer intentions and behaviors toward ethical products. *Journal of Marketing*, 76(1), 103-118

APPENDIX A – Experimental Instructions (translated from Italian).

Instructions

Welcome! Thank you for participating in this experiment.

For your participation you will receive €3. On top of this an additional sum may be paid to you depending on the decisions made – by you and the other participants – during the experiment.

During the experiment it is prohibited to talk with other participants: if you have any question at any time please raise your hand and wait for an experimenter to come to your desk and answer it in private. We kindly ask you to turn off the cell phone. In case you violate these rules, you will be asked to leave the experiment and you will lose the right to be paid.

After reading the instructions, to ensure the full comprehension of the experiment, we will ask you to complete a questionnaire through the computer that has been assigned to you. The experiment will begin after all participants answered correctly all questions.

During the experiment, all information and decisions made will be anonymous and confidential. At the end of the experiment each of you will be paid in private and in cash.

Experimental instructions

In this experiment there are two roles. Each participant will be identified either as Player A or Player B. Roles are assigned randomly and will remain the same for the whole experiment.

Each participant has an initial endowment in Euro that varies based on the role:

- **Player A** has an initial endowment of €9.
- **Player B** has an initial endowment of €0.

You will be informed if you are Player A or B at the beginning of the experiment.

For each pair of players (formed by one Player A and one Player B), an initial donation of €5 has been allocated to the organization *Doctors without borders*. The final amount of the donation depends on the decisions made in the experiment. For those who don't know it, *Doctors without borders* is described on Wikipedia as follows:

"*Doctors Without Borders is a private international organization whose purpose is to bring emergency aid and health care assistance in areas of the world where the right to health care is not yet guaranteed.*"

Doctors without borders asks to support, through voluntary donations, the independent and effective intervention in over 70 countries around the world, to fight diseases such as malnutrition and malaria, to fulfill vaccination programs and to guarantee health care assistance saving every day human lives.

See below a screenshot of the webpage through which the organization asks for donations.

Interaction rules between Player A and Player B

The experiment is divided in three phases.

PHASE 1

In this phase, each Player A has the opportunity to increase their own endowment by taking money from the donation to *Doctors without borders*. Each Player A can take between 0 and 5 Euros in intervals of 1 Euro. Thus, they can take 0, 1, 2, 3, 4 or 5 Euros. What is not taken by Player A (remaining donation) will be donated to *Doctors without borders*.

At the end of PHASE 1 each Player A will have their initial endowment of €9 plus the amount taken from *Doctors without borders (if any)*.

PHASE 2

In this phase, each Player A has to make an offer on how to split the amount they have to the Player B they will be matched with. Player A can choose between two alternatives:

- Offer €**5.**
- Offer **€1.**

PHASE 3

In this phase, each Player B will have to make a sequence of decisions regarding the offers made by Players A in PHASE 2.

[NO-INFO: For every Player A, Player B will be informed of:

- The amount offered (€5 or €1) in PHASE 2.

Note: Player B will never know (even at the end of today's experiment), how much each Player A has taken from *Doctors without borders*.***]***

[INFO & INFO-RESTORE: For every Player A, Player B will be informed of:

- The amount taken from *Doctors without borders* in PHASE 1.
- The amount offered (€5 or €1) in PHASE 2.***]***

For every offer, Player B has to decide whether to **accept** or **reject** it:

- If Player B **accepts**, the offer proposed is implemented.
- If Player B **rejects**, both Player A and Player B receive €0.
  [INFO-RESTORE: If Player B **rejects**, both Player A and Player B receive €0, while *Doctors without Borders* receives the initial donation of €5***]***.

Matching process

Each Player B is actually matched with only **one** Player A.

However, Player B will not know, until the end of the experiment, which Player A they are assigned to. Thus, they will not know which of the decisions described above will be relevant for their earnings.

Therefore, the best strategy is to treat all the decisions as relevant as any of them could be relevant.

If you have any question about the experiment, please raise your hand and wait for an experimenter to come to your desk and answer it in private. If there are no questions we now proceed with the questionnaire and, later, with the experiment.

APPENDIX B – Alternative estimations to Table 4

Table B1. Mixed-effects probit.

	All offers (1)	€1 offers (2)	€5 offers (3)
INFO-RESTORE	-0.10 (0.27)	-0.23 (0.27)	-1.71** (0.68)
Amount taken	-0.43*** (0.29)	-0.19* (0.11)	-0.95*** (0.18)
Constant	1.77*** (0.27)	-0.16 (0.35)	5.45*** (0.97)
N	442	126	316

Notes: mixed-effects probit with random intercepts only at session level. Dependent variable takes value 1 if the offer is accepted and zero otherwise. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table B2. Mixed-effects linear model.

	All offers (1)	€1 offers (2)	€5 offers (3)
INFO-RESTORE	-0.03 (0.09)	-0.03 (0.09)	-0.13* (0.07)
Amount taken	-0.14*** (0.02)	-0.04** (0.02)	-0.12*** (0.02)
Constant	1.09*** (0.08)	0.35*** (0.09)	1.26*** (0.06)
N	442	126	316

Notes: mixed-effects linear model with random intercepts at individual and session level. Dependent variable takes value 1 if the offer is accepted and zero otherwise. * significant at 10%, ** significant at 5%, *** significant at 1%.

APPENDIX C – Behavioral predictions

C.1 Inequity aversion

For the responder (R henceforth), the utility has the following functional form

$$U_R(\pi_P, \pi_R, \pi_{NGO}) = \pi_R - \frac{\alpha}{2}[\max(\pi_P - \pi_R\,, 0) + \max(\pi_{NGO} - \pi_R\,, 0)] - \frac{\beta}{2}[\max(\pi_R - \pi_P\,, 0) + \max(\pi_R - \pi_{NGO}\,, 0)] \quad (1)$$

where,

- $\pi_P = E_P + t - o$, if R accepts ($\pi_P = 0$ otherwise)
- $\pi_R = E_R + o$, if R accepts ($\pi_R = 0$ otherwise)
- $\pi_{NGO} = E_{NGO} - t$, if R accepts; $\pi_{NGO} = E_{NGO} - t$, if R rejects in INFO; $\pi_{NGO} = E_{NGO}$, if R rejects in INFO-RESTORE.
- $E_P = 9, E_R = 0$, and $E_{NGO} = 5$.

and by assumption from Fehr and Schmidt (1999),

- $0 \leq \beta < 1$
- $\alpha \geq \beta$.

C.1.a. Proposer offers 5 ($o = 5$)

When the proposer (P henceforth) offers 5, the responder compares acceptance and rejection. If R accepts their utility is equal to

$$U_R(t) = 5 - \frac{\alpha}{2}[\max(t - 1\,, 0) + \max(-t\,, 0)] - \frac{\beta}{2}[max(1 - t\,, 0) + max(t\,, 0)].$$

Given the discontinuity of the utility function conditional on t, we can rewrite it as

$$U_R(t) = \begin{cases} 5 - \frac{\beta}{2} & if\ t \leq 1 \quad (2) \\ 5 - \frac{\alpha}{2}(t-1) - \frac{\beta}{2}t & if\ t > 1 \quad (3). \end{cases}$$

When R rejects, in INFO their utility as a function of the amount taken t has the following functional form:

$$U_R(t) = 0 - \frac{\alpha}{2}[\max(0\,, 0) + \max(5 - t\,, 0)] - \frac{\beta}{2}[max(0\,, 0) + max(-5 + t\,, 0)]$$

And since $5 - t \geq 0, \forall t$ and $-5 + t \leq 0, \forall t$; it derives that

$$U_R(t) = -\frac{\alpha}{2}(5 - t) \quad (4).$$

Finally, in INFO-RESTORE, when R rejects her utility is constant and equal to

$$U_R = -\frac{5}{2}\alpha \quad (5).$$

In INFO, R accepts P's offer if and only if either $(2) > (4)$ (for $t \leq 1$) or $(3) > (4)$ (for $t > 1$). Similarly, in INFO-RESTORE, R accepts P's offer if and only if either $(2) > (5)$ (for $t \leq 1$) or $(3) > (5)$ (for $t > 1$). In what follows, we analyze these cases in turn.

INFO treatment

In INFO and for $t \leq 1$, R accepts if and only if the equation (2) is greater than (4), which leads to the following condition:

$$t < \frac{10 - \beta + 5\alpha}{\alpha}.$$

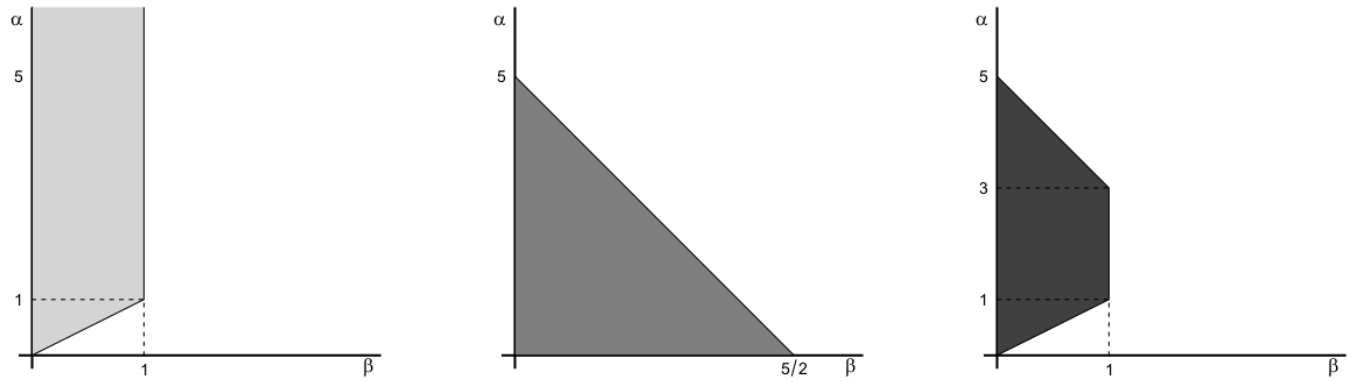
Notice that given the assumptions on α and β, the above inequality holds for any combinations of α and β, for both $t = 0$ and $t = 1$.[9] This implies that, in INFO, R will always accept an offer of 5 when P has taken 0 or 1 from the NGO endowment.

For $t > 1$, R accepts if and only if equation (3) is greater than equation (4), which leads to the following condition:

$$t < \frac{10 + 6\alpha}{2\alpha + \beta}.$$

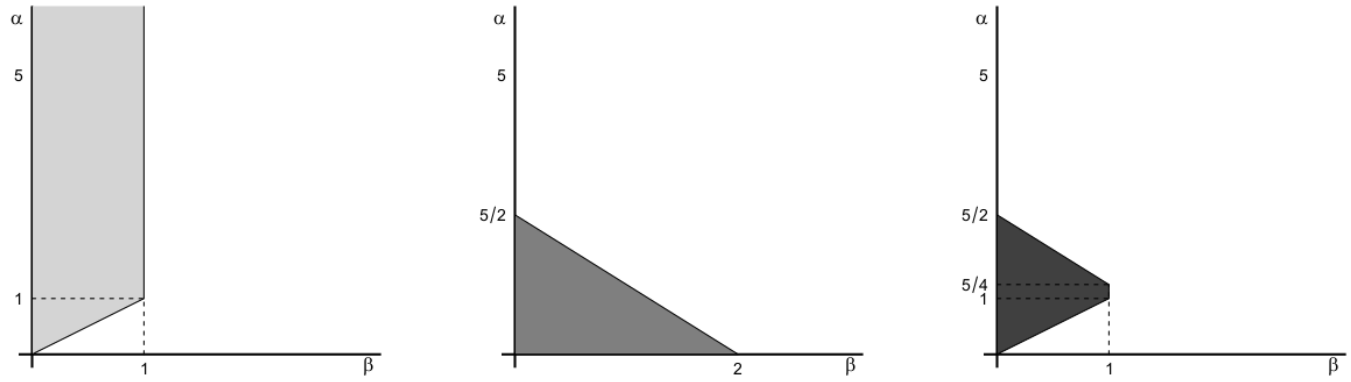
As before, given the assumptions on α and β, the inequality always holds for $t = 2$ and $t = 3$ but it holds only for some combinations of α and β when P takes 4 or 5. Thus, in these cases, R will not always accept. Graphically the combinations of α and β that lead R to accept are represented by the dark grey area in Panel C of Figure C1 (for $t = 4$) and Figure C2 (for $t = 5$). These areas are the result of the intersection of the areas in Panel A (in light grey, which represents the conditions $0 \leq \beta < 1$; $\alpha \geq \beta$) and Panel B (in grey, which represents $t < \frac{10+6\alpha}{2\alpha+\beta}$ for $t = 4$ and $t = 5$, respectively). Only those combinations of α and β in the dark grey area satisfy at the same time both the conditions imposed by the assumptions and the inequality. From the comparison of Panels C between Figures C1 and C2, it is easy to see that the likelihood of acceptance decreases with the increase in the amount taken.

[9] Here and in the rest of the analysis, we exclude the case $\alpha = \beta = 0$, in which the utility function becomes the standard case with no concerns for inequity.



Panel A Panel B Panel C

Figure C1. INFO - R's acceptance area when $t = 4$ and $o = 5$.



Panel A Panel B Panel C

Figure C2. INFO - R's acceptance area when $t = 5$ and $o = 5$.

INFO-RESTORE treatment

In INFO-RESTORE and for $t \leq 1$, R accepts if and only if equation (2) is greater than equation (5), which leads to the following condition:

$$\alpha > \frac{\beta - 10}{5}.$$

Given our assumptions on α and β, the above inequality holds for any value of α and β. Thus, as in the INFO treatment, the responder will accept an offer of 5 both when $t = 0$ and $t = 1$.

For $t > 1$, R accepts if and only if

$$t < \frac{10 + 6\alpha}{\alpha + \beta}.$$

Also in this case, given the assumption on α and β, the inequality holds $\forall t > 1$ for any combination of α and β.

To summarize, while in INFO, for some values of α and β, offers of 5€ are rejected this is never the case in INFO-RESTORE, leading to the treatment prediction of PROPOSITION 1a in the main text.

C.1.b. Proposer offers 1 ($o = 1$)

When P offers 1 and R accepts her utility is equal to

$$U_R(t) = 1 - \frac{\alpha}{2}[\max(7 + t\,, 0) + \max(4 - t\,, 0)] - \frac{\beta}{2}[max(-7 - t\,, 0) + max(t - 4\,, 0)].$$

Given restrictions on t, α, and β we can re-write the utility of acceptance as:

$$U_R = \begin{cases} 1 - \frac{11}{2}\alpha & if\ t \leq 4 \quad (6) \\ 1 - 6\alpha - \frac{\beta}{2} & if\ t = 5 \quad (7). \end{cases}$$

In case of rejections, nothing changes from the case of $o = 5$. Thus, in INFO, the responder utility function is equal to equation (4), while in INFO-RESTORE it is equal to equation (5). Hence, in INFO, R accepts P's offer if and only if either $(6) > (4)$ (for $t \leq 4$) or $(7) > (4)$ (for $t = 5$). Similarly, in INFO-RESTORE, R accepts P's offer if and only if either $(6) > (5)$ (for $t \leq 4$) or $(7) > (5)$ (for $t = 5$). In what follows, we analyze these cases in turn.

INFO treatment

In INFO and for $t \leq 4$, R accepts if and only if equation (6) is greater than equation (4), which leads to the following condition:

$$t < \frac{2 - 6\alpha}{\alpha}.$$

The above inequality holds, according to the amount taken, only for specific values of α, which represent the threshold for R's acceptance. These values are summarized in Table C1.

Table C1. INFO - R's acceptance thresholds when $t \leq 4$ and $o = 1$

Amount taken (*t*)	Condition for R's acceptance
0	$0 \leq \beta \leq \alpha < \frac{1}{3}$
1	$0 \leq \beta \leq \alpha < \frac{2}{7}$
2	$0 \leq \beta \leq \alpha < \frac{1}{4}$
3	$0 \leq \beta \leq \alpha < \frac{2}{9}$
4	$0 \leq \beta \leq \alpha < \frac{1}{5}$

In INFO and for $t = 5$, R accepts if and only if equation (7) is greater than equation (4), which leads to the following condition:

$$\alpha < \frac{2-\beta}{12}.$$

Notice that $\frac{2-\beta}{12} < \frac{1}{5} \; \forall \; 0 \leq \beta < 1$. In other words, R's acceptance threshold when $t = 5$ is smaller than the one when $t = 4$.

In sum, acceptance in INFO decreases as t increases.

INFO-RESTORE treatment

In INFO-RESTORE and for $t \leq 4$, R accepts if and only if equation (6) is greater than equation (5), which leads to the following condition:

$$\alpha < \frac{1}{3}.$$

For $t = 5$, R accepts if and only if equation (7) is greater than equation (5), which leads to the following condition:

$$\alpha < \frac{2-\beta}{7}.$$

Notice that $\frac{2-\beta}{7} < \frac{1}{3} \; \forall \; 0 \leq \beta < 1$. In other words, R's acceptance thresholds when $t = 5$ is smaller than the ones for $t \leq 4$.

Comparing the acceptance thresholds between INFO and INFO-RESTORE, we find that also for offers of 1€ acceptance is more likely in INFO-RESTORE than in INFO for $t > 0$. This proves PROPOSITION 1b in the main text.

C.1.c Summary of predictions according inequity aversion

Table C2 summarizes predicted differences in acceptance between INFO and INFO-RESTORE for all possible levels of taking. Given these predicted difference, we expect according to Fehr and Schmidt (1999) higher acceptance rates in INFO-RESTORE than in INFO.

Table C2. Difference in acceptance between INFO and INFO-RESTORE.

	Amount taken					
Offer	*0*	*1*	*2*	*3*	*4*	*5*
$o = 5$	=	=	=	=	−	−
$o = 1$	=	−	−	−	−	−

Notes: = means that acceptance in INFO-RESTORE is as likely as it is in INFO. + means that acceptance in INFO is more likely than in INFO-RESTORE. − means that acceptance in INFO-RESTORE is more likely than in INFO.

C.2 Inequity aversion toward the Proposer and consequentialist morality toward the NGO

We assume that R cares about inequity toward P but they suffer disutility if the NGO receives less than the maximal donation. Their utility has the following functional form:

$$U_R(\pi_P, \pi_R, \pi_{NGO}) = \pi_R - \alpha[\max(\pi_P - \pi_R\,, 0)] - \beta[\max(\pi_R - \pi_P\,, 0)] - \gamma(5 - \pi_{NGO}) \qquad (8)$$

where,

- $\pi_P = E_P + t - o$, if R accepts ($\pi_P = 0$ otherwise)
- $\pi_R = E_R + o$, if R accepts ($\pi_R = 0$ otherwise)
- $\pi_{NGO} = E_{NGO} - t$, if R accepts; $\pi_{NGO} = E_{NGO} - t$, if R rejects in INFO; $\pi_{NGO} = E_{NGO}$, if R rejects in INFO-RESTORE.
- $E_P = 9, E_R = 0$, and $E_{NGO} = 5$.

and by assumption,

- $0 \leq \beta < 1$
- $\alpha \geq \beta$
- $0 < \gamma < 1$.

C.2.a. Proposer offers 5 ($o = 5$)

When P offers 5 and R accepts the offer, their utility is equal to

$$U_R(t) = 5 - \alpha[\max(t - 1\,, 0)] - \beta[max(1 - t\,, 0)] - \gamma t.$$

Given restrictions on t, α, β, and γ we can re-write the utility of acceptance as:

$$U_R(t) = \begin{cases} 5 - \beta & if\ \ t = 0 \qquad (9) \\ 5 - \gamma & if\ t = 1 \qquad (10) \\ 5 - \alpha(t - 1) - \gamma t & if\ t > 1 \qquad (11). \end{cases}$$

In INFO, when R rejects, her utility as a function of the amount taken (t) has the following form

$$U_R(t) = 0 - \alpha[\max(0\,, 0)] - \beta[max(0\,, 0)] - \gamma t$$

and it derives that

$$U_R(t) = \begin{cases} 0 & if \quad t = 0 \qquad (12) \\ -\gamma t & if \quad t > 0 \qquad (13). \end{cases}$$

Finally, in INFO-RESTORE, when R rejects her utility becomes constant and equal to zero:

$$U_R(t) = 0 - \alpha[\max(0\,,0)] - \beta[max(0\,,0)] - 0\gamma = 0 \qquad (14).$$

It follows that, in INFO, R accepts P's offer if and only if $(9) > (12)$ (for $t = 0$), $(10) > (13)$ (for $t = 1$), or $(11) > (13)$ (for $t > 1$). Similarly, in INFO-RESTORE, R accepts P's offer if and only if $(9) > (14)$ (for $t = 0$), $(10) > (14)$ (for $t = 1$), or $(11) > (14)$ (for $t > 1$). We analyze all these cases in turn.

INFO treatment

In INFO and for $t = 0$, R accepts if and only if $(9) > (12)$, which implies:

$$5 - \beta > 0.$$

Given our assumptions on β, the above inequality always holds. Thus, R will accept an offer of 5 when $t = 0$.

For $t = 1$, R accepts if and only if $(10) > (13)$, implying:

$$5 - \gamma > -\gamma t.$$

Again, the above inequality always holds. Thus, R will accept an offer of 5 when $t = 1$.

Finally, for $t > 1$, R accepts if and only if $(11) > (13)$, which leads to the following condition:

$$t < \frac{5 + \alpha}{\alpha}.$$

The above inequality holds, according to the amount taken, only for specific values of α, which represent the threshold for R's acceptance. These values are summarized in Table C3.

Table C3. INFO - R's acceptance thresholds when $t > 1$ and $o = 5$

Amount taken by P	Condition for R's acceptance
2	$0 \leq \beta \leq \alpha < 5$
3	$0 \leq \beta \leq \alpha < \frac{5}{2}$
4	$0 \leq \beta \leq \alpha < \frac{5}{3}$
5	$0 \leq \beta \leq \alpha < \frac{5}{4}$

Thus it follows that, the less P takes the higher is R's acceptance threshold. Hence, in this respect this model offers similar predictions to the standard Fehr and Schmidt (1999) model as acceptance is predicted to decrease as t increases.

INFO-RESTORE treatment

In INFO-RESTORE and for $t = 0$, R accepts if and only if $(9) > (14)$, implying:

$$5 - \beta > 0.$$

Given our assumptions on β, the above inequality always holds. Thus, similar to the INFO case, R will accept an offer of 5 when $t = 0$.

For $t = 1$, R accepts if and only if $(10) > (14)$, implying:

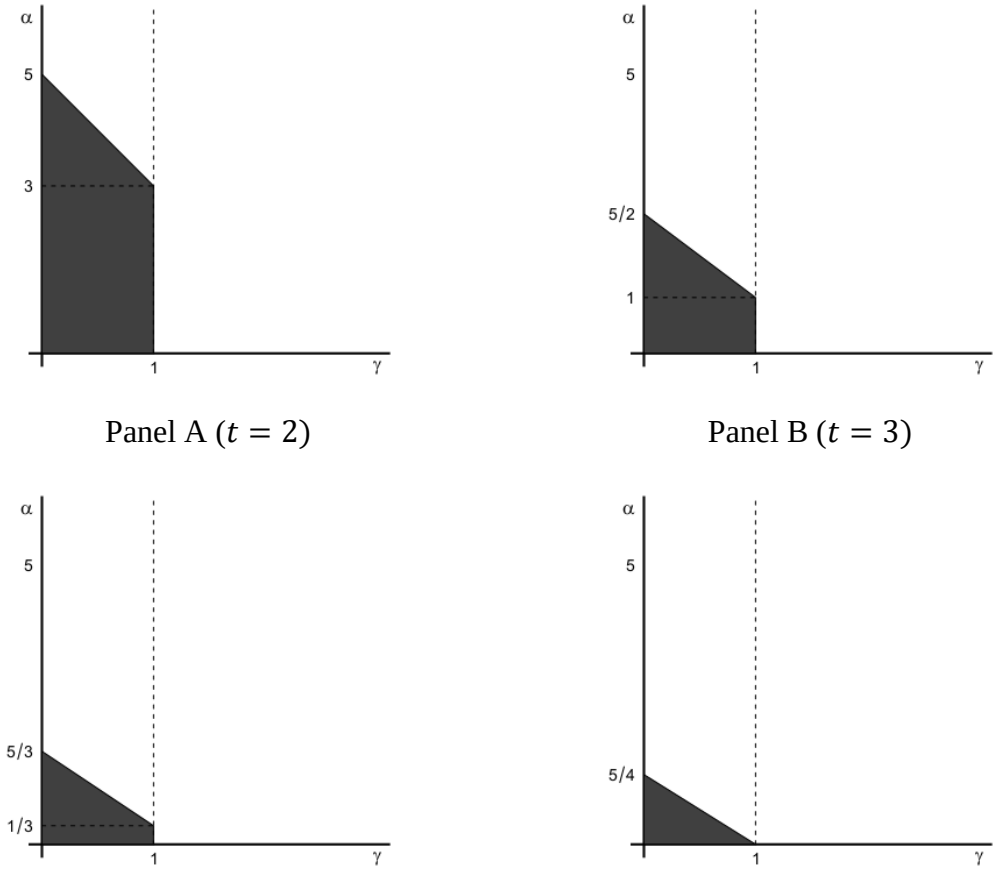
$$5 - \gamma > 0.$$

Given our assumptions on γ, the above inequality always holds. Thus, similar to the INFO case, R will accept an offer of 5 when $t = 1$.

Finally, for $t > 1$, R accepts if and only if $(11) > (14)$, implying:

$$t < \frac{5 + \alpha}{\alpha + \gamma}.$$

Given our assumptions on α and γ, the above inequality holds only for specific combinations of the two parameters. Figure C3 summarizes the possible combinations of α and γ that lead to R's acceptance for different values of t. As it is clear from Figure C3, we should expect less acceptance as t increases as in the INFO case.



Panel A ($t = 2$) Panel B ($t = 3$)

Panel C ($t = 4$) Panel D ($t = 5$)

Figure C3. INFO - R's acceptance area when $t > 1$ and $o = 5$.

Finally, given that $\frac{5+\alpha}{\alpha+\gamma} < \frac{5+\alpha}{\alpha}$ always holds for positive values of γ, it follows that the area of acceptance for a given *t* is always greater in INFO than in INFO-RESTORE, summarized in PROPOSITION 2a in the main text. Notice that this effect goes in opposite direction compared to inequity aversion.

C.2.b. Proposer offers 1 ($o = 1$)

When P offers 1, if R accepts their utility is equal to

$$U_R(t) = 1 - \alpha[\max(7 + t\,, 0)] - \beta[max(-7 - t\,, 0)] - \gamma t\,.$$

Since for any t, 7+$t > 0$ and $-7 - t < 0$, R's utility results to be:

$$U_R(t) = 1 - \alpha(7 + t) - \gamma t \qquad (15).$$

Nothing changes, in case of rejections, from the case of $o = 5$. Thus, in INFO, R's utility function is equal to either (12) or (13), while in INFO-RESTORE it is equal to (14).

INFO treatment

In INFO and for $t = 0$, R accepts the offer if and only if $(15) > (12)$, implying:

$$1 - 7\alpha > 0$$

Thus, R will accept an offer of 1 when $t = 0$ only if $\alpha < \frac{1}{7}$.

For $t > 0$, R accepts the offer if and only if $(15) > (13)$, implying:

$$t < \frac{1 - 7\alpha}{\alpha}.$$

The above inequality holds, according to the amount taken, only for specific values of α, which represent the threshold for R's acceptance. These values are summarized in Table C4.

Table C4. INFO - R's acceptance thresholds when t > 0 and *o* = 1.

Amount taken by P	Condition for R's acceptance
1	$0 \leq \beta \leq \alpha < \frac{1}{8}$
2	$0 \leq \beta \leq \alpha < \frac{1}{9}$
3	$0 \leq \beta \leq \alpha < \frac{1}{10}$
4	$0 \leq \beta \leq \alpha < \frac{1}{11}$
5	$0 \leq \beta \leq \alpha < \frac{1}{12}$

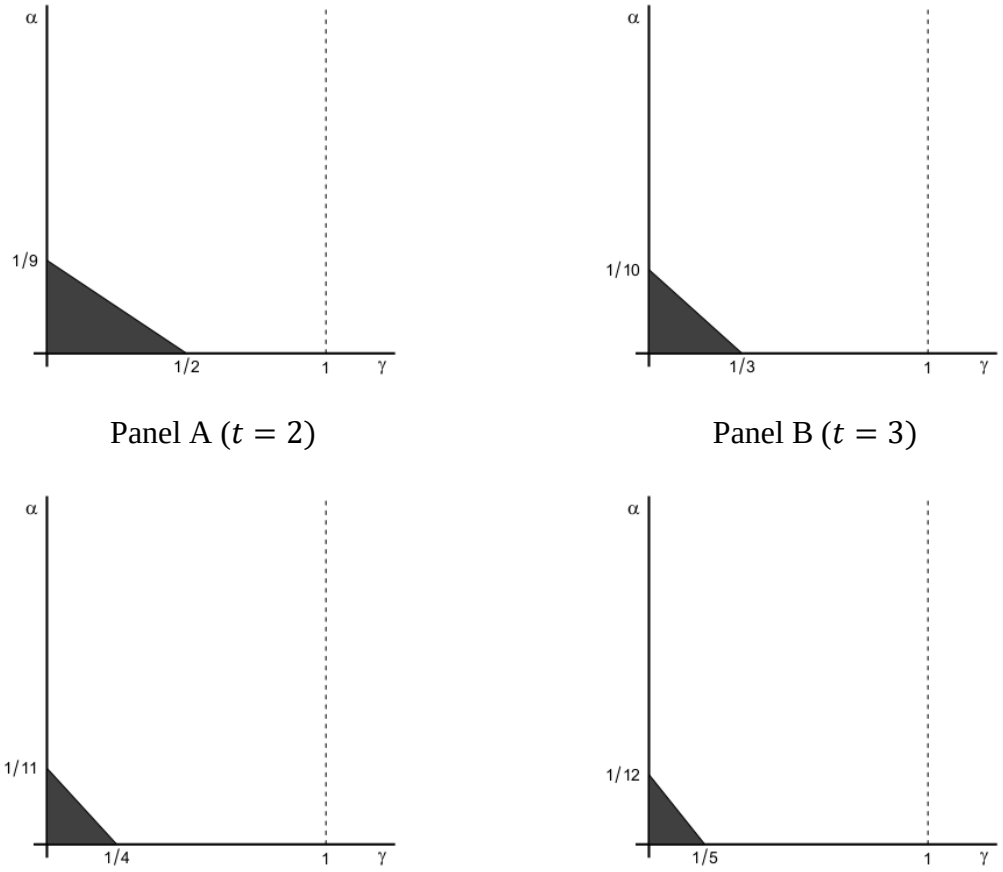
Thus, it follows that, the less P takes the higher is R's acceptance threshold: acceptance happens for very low levels of α and it is less likely when t increases.

INFO-RESTORE treatment

In INFO-RESTORE for any t, R accepts the offer if and only if $(15) > (14)$, implying:

$$t < \frac{1 - 7\alpha}{\alpha + \gamma}.$$

Given our assumptions on α and γ, the above inequality holds only for specific combinations of the two parameters. Figure C4 summarizes the possible combinations of α and γ that lead to R's acceptance for different values of t. As it is clear from Figure C4, also in INFO-RESTORE we should expect less acceptance as t increases.



Panel A ($t = 2$) Panel B ($t = 3$)

Panel C ($t = 4$) Panel D ($t = 5$)

Figure C4. INFO - R's acceptance area when $t > 1$ and $o = 1$.

Similar to the case of offers of 5€, to compare acceptance across the two treatments we compare the acceptance thresholds and since $\frac{1-7\alpha}{\alpha+\gamma} < \frac{1-7\alpha}{\alpha}$, we conclude that acceptance areas are larger in INFO than in INFO-RESTORE. This is summarized in PROPOSITION 2b in the main text. As for offers of 5€, this prediction goes in the opposite direction compared to inequity aversion.

C.2.c Summary of predictions according to consequentialist morality

Table C5 summarizes predicted differences in acceptance between INFO and INFO-RESTORE for all possible levels of taking. Given these predicted difference, we expect higher acceptance rates in INFO than in INFO-RESTORE.

Table C5. Difference in acceptance between INFO and INFO-RESTORE.

	Amount taken					
Offer	*0*	*1*	*2*	*3*	*4*	*5*
$o = 5$	=	=	+	+	+	+
$o = 1$	=	+	+	+	+	+

Notes: = means that acceptance in INFO-RESTORE is as likely as it is in INFO. + means that acceptance in INFO is more likely than in INFO-RESTORE. − means that acceptance in INFO-RESTORE is more likely than in INFO.

C.3 Inequity aversion toward the Proposer and Kantian morality toward the NGO

We assume that R cares about inequity toward P but they suffer a cost whenever they accept an offer from the proposer and a positive amount has been taken from the NGO. Their utility has the following functional form:

$$U_R(\pi_P, \pi_R, \pi_{NGO}) = \pi_R - \alpha[\max(\pi_P - \pi_R\ , 0)] - \beta[\max(\pi_R - \pi_P\ , 0)] - \mathbb{1}_{t>0}\mathbb{1}_{Accept}K$$

where,

- $\pi_P = E_P + t - o$, if R accepts ($\pi_P = 0$ otherwise)
- $\pi_R = E_R + o$, if R accepts ($\pi_R = 0$ otherwise)
- $\pi_{NGO} = E_{NGO} - t$, if R accepts; $\pi_{NGO} = E_{NGO} - t$, if R rejects in INFO; $\pi_{NGO} = E_{NGO}$, if R rejects in INFO-RESTORE.
- $E_P = 9, E_R = 0$, and $E_{NGO} = 5$.

and by assumption,

- $0 \leq \beta < 1$
- $\alpha \geq \beta$
- $K > 0$.

C.3.a. Proposer offers 5 ($o = 5$)

When P offers 5 and R accepts the offer, their utility is equal to

$$U_R(t) = 5 - \alpha[\max(t - 1\ , 0)] - \beta[max(1 - t\ , 0)] - \mathbb{1}_{t>0}K.$$

Given restrictions on t, α, and β we can re-write the utility of acceptance as:

$$U_R(t) = \begin{cases} 5 - \beta & if\ \ t = 0 \quad (16) \\ 5 - K & if\ t = 1 \quad (17) \\ 5 - \alpha(t-1) - K & if\ t > 1 \quad (18). \end{cases}$$

Both in INFO and in INFO-RESTORE, when R rejects, her utility has the following form

$$U_R = 0 \qquad (19)$$

given that the responder only suffers a moral cost in case if they accept the offer (and the offer is 'immoral').

INFO and INFO-RESTORE treatment

In both INFO and INFO-RESTORE and for $t = 0$, R accepts if and only if $(16) > (19)$, which implies:

$$5 - \beta > 0.$$

Given our assumptions on β, the above inequality always holds. Thus, R will accept an offer of 5 when $t = 0$.

For $t = 1$, R accepts if and only if $(17) > (19)$, implying:

$$5 - K > 0.$$

The above inequality holds depending on the moral cost K. Thus, R will accept an offer of 5 when $t = 1$ only if $K < 5$.

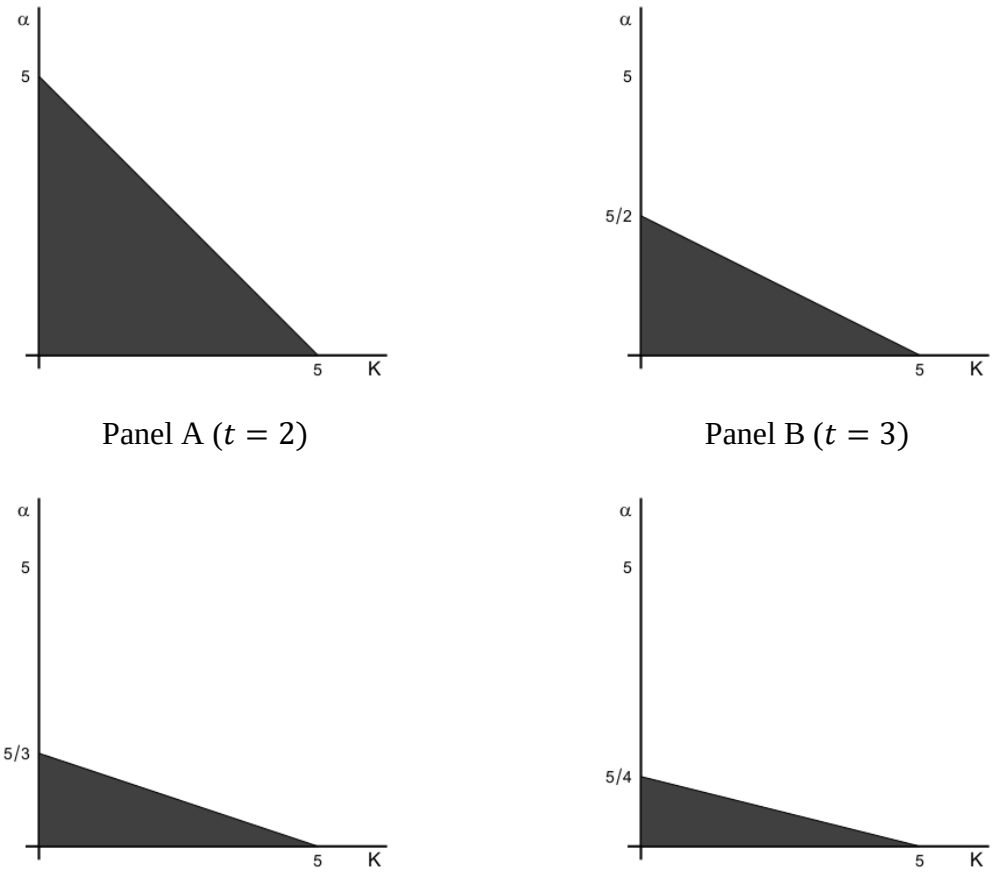
Finally, for $t > 1$, R accepts if and only if $(18) > (19)$, which leads to the following condition:

$$t < \frac{5 + \alpha - K}{\alpha}.$$

The above inequality holds, according to the amount taken, only for specific values of α and K, which represent the threshold for R's acceptance. The combinations of α and K that lead to acceptance are summarized in Figure C5.

As it is clear from the comparisons of areas in panels in Figure C5, the less P takes the higher is R's acceptance threshold. Hence, in this respect this model offers similar predictions to the two models previously considered.

However, it is worth noting that, in contrast to the other two models, predictions on acceptance are in this case identical in INFO and INFO-RESTORE since the responder's moral cost does not depend on the payoff consequences on the NGO.



Panel A ($t = 2$) Panel B ($t = 3$)

Panel C ($t = 4$) Panel D ($t = 5$)

Figure C5. R's acceptance area when $t > 1$ and $o = 5$.

C.3.b. Proposer offers 1 ($o = 1$)

When P offers 1, if R accepts their utility is equal to

$$U_R(t) = 1 - \alpha[\max(7 + t\,, 0)] - \beta[max(-7 - t\,, 0)] - \mathbb{1}_{t>0}K.$$

Since for any t, 7+$t > 0$ and $-7 - t < 0$, R's utility for $t = 0$ is:

$$U_R(t) = 1 - 7\alpha \qquad (20).$$

while R's utility for $t > 0$ is:

$$U_R(t) = 1 - \alpha(7 + t) - K \qquad (21).$$

As in the case of offers of 5, both in INFO and in INFO-RESTORE, when R rejects, her utility is equal to

$$U_R = 0 \qquad (22)$$

given that the responder only suffers a moral cost in case of acceptance.

INFO and INFO-RESTORE treatment

Both in INFO and INFO-RESTORE for $t = 0$, R accepts the offer if and only if $(20) > (22)$, implying:

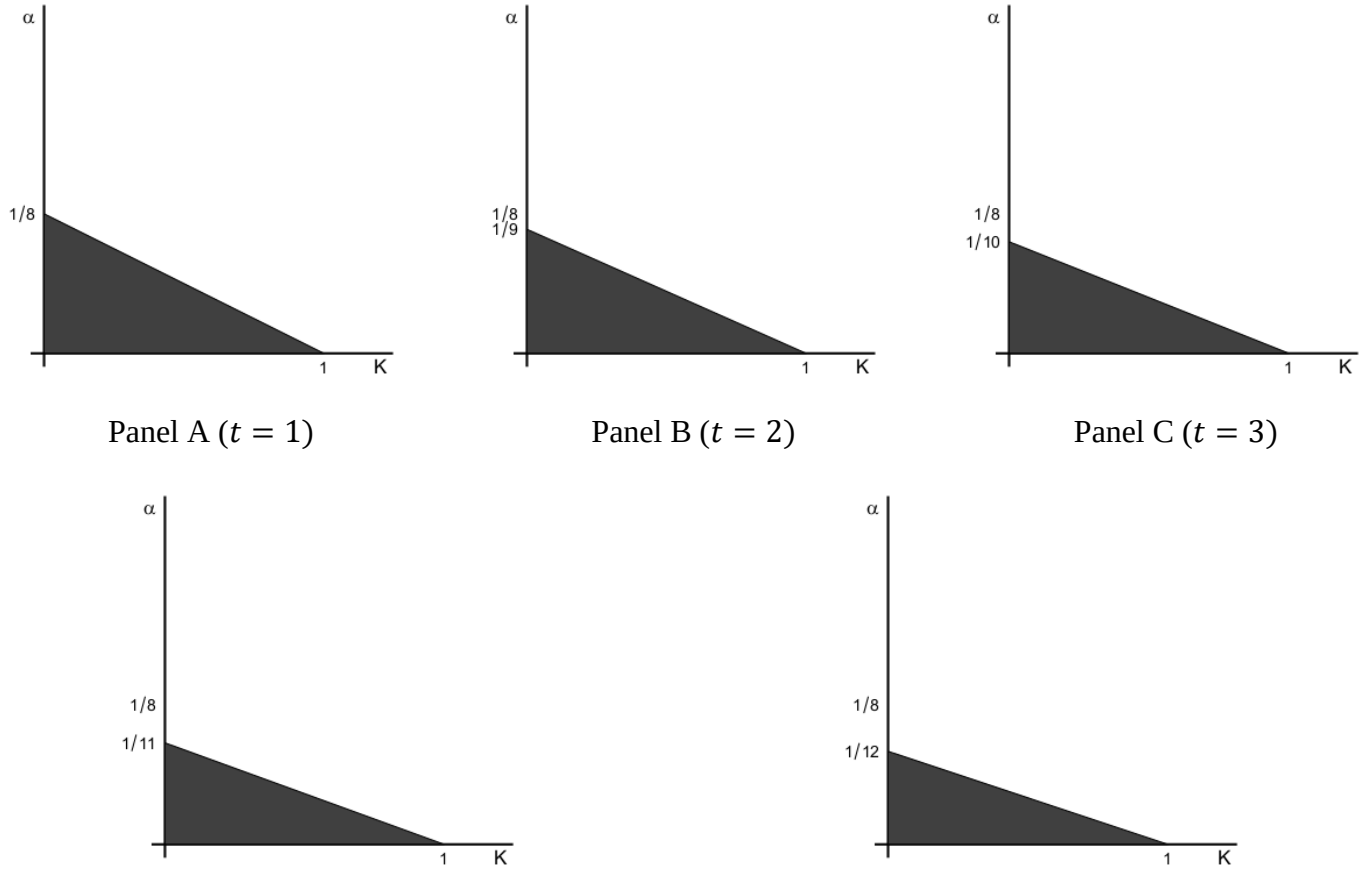
$$1 - 7\alpha > 0.$$

Thus, R will accept an offer of 1 when $t = 0$ only if $\alpha < \frac{1}{7}$.

For $t > 0$, R accepts the offer if and only if $(21) > (22)$, implying:

$$t < \frac{1 - 7\alpha - K}{\alpha}.$$

Figure C6 summarizes and compares R's acceptance thresholds. As for the case of $o = 5$, also in the case of $o = 1$ acceptance is less likely as the amount taken increases. Also in this case, predictions on acceptance do not differ across INFO and INFO-RESTORE.



Panel A ($t = 1$) Panel B ($t = 2$) Panel C ($t = 3$)

Panel D ($t = 4$) Panel E ($t = 5$)

Figure C6. R's acceptance area when $t > 0$ and $o = 1$.

C.3.c Summary of predictions according to Kantian morality

To sum up, the Kantian model predicts acceptance to decrease with an increase of the amount taken by the proposer, but it predicts no difference between INFO and INFO-RESTORE because the moral cost is not triggered by the consequences for the NGO, but rather by own actions.