Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Frustration and Anger in the Ultimatum Game: An Experiment

Chiara Aina, Pierpaolo Battigalli, Astrid Gamba

Working Paper n. 621

This Version: July, 2019

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

The opinions expressed in the working papers are those of the authors alone, and not those of the Institute, which takes non institutional policy position, nor those of CEPR, NBER or Università Bocconi.
Frustration and Anger in the Ultimatum Game: An Experiment

Chiara Aina∗  Pierpaolo Battigalli†  Astrid Gamba‡

Abstract

In social dilemmas, choices may depend on belief-dependent motivations enhancing the credibility of promises or threats at odds with personal gain maximization. We address this issue theoretically and experimentally in the context of the Ultimatum Minigame, assuming that the choice of accepting or rejecting an unfair proposal is affected by a combination of frustration, due to unfulfilled expectations, and inequity aversion. We increase the responder’s payoff from the default allocation (the proposer’s outside option) with the purpose of increasing the responder’s frustration due to the unfair proposal, and thus his willingness to reject it. In addition, we manipulate the method of play, with the purpose of switching on (direct response method) and off (strategy method) the responder’s experience of anger. Our behavioral predictions across and within treatments are derived from the theoretical model complemented by explicit auxiliary assumptions, without relying on equilibrium analysis.

Keywords: Experiments, psychological games, ultimatum minigame, frustration, anger, non-equilibrium analysis.

JEL classification: C72, C91, D03.

∗University of Zurich. Schönberggasse 1, 8001 Zürich, Switzerland; e-mail: chiara.aina@econ.uzh.ch.
†Corresponding author. Bocconi University and IGIER. Via Guglielmo Röntgen, 1, 20136 Milano, Italy; e-mail: pierpaolo.battigalli@unibocconi.it.
‡Università degli Studi dell’Insubria. Via Monte Generoso 71, 21100 Varese, Italy; e-mail: astrid.gamba@uninsubria.it.
1 Introduction

In this paper we test experimentally the theory of frustration and anger by Battigalli, Dufwenberg and Smith (2019)—henceforth BDS—in the strategic context of the Ultimatum Minigame. In this social dilemma, a proposer can decide to make a fair offer that is automatically accepted, or a greedy offer that the responder can either accept or reject. Since the fair offer is automatically accepted, we call this choice of the proposer the “default allocation.” Experimental results about behavior in this and in similar negotiation scenarios (e.g., Ultimatum Game) systematically deviated from the predictions of traditional economic theory, reporting a positive frequency of both default allocations and rejections of greedy offers. We ask whether the responder’s choice of rejecting a greedy offer can be explained by belief-dependent preferences in line with BDS’s theory of anger from blaming behavior. Upon observing a greedy offer, the responder may experience frustration, which is measured as the gap between his initially expected payoff and the maximal payoff that he can achieve after the greedy offer. Frustration triggers anger that may result in a rejection, depending on the responder’s sensitivity to this emotion.

In the experiment, subjects play one-shot and anonymously an Ultimatum Minigame (henceforth UMG). Our experimental manipulation is twofold, affecting both the payoff structure and the method of play in a $2 \times 2$ design. We manipulate the responder’s monetary payoff from the default allocation to increase his initial expectations and thus his frustration in case of a greedy offer. In addition, we manipulate the method of play, i.e., direct response versus strategy method, to determine whether responders actually observe the greedy offer before making their choice and can thus experience frustration. Indeed, when the game is played with the strategy method, responders have to decide whether to accept or reject a greedy offer without observing whether it was actually made. It is rather implausible that, by only contemplating the possibility of facing a greedy offer, responders feel frustrated as if they had actually received it. In fact, it is one thing to imagine ourselves being frustrated due to an hypothetical event, and quite another thing to actually be frustrated due to an accomplished fact. Thus, anger should have no bite with

---

1 A large number of studies, conducted with different incentives and in different countries, found a robust pattern in the Ultimatum Game: the majority of proposers offer 40% to 50% of the total sum (with the modal offer being an even split), while responders reject offers below 20% (Güth et al. 1982, Thaler 1988, Güth and Tietz 1990, Bolton and Zwick 1995).

2 Notice that, for all leader-follower games (like the Ultimatum Game), BDS’s model of anger from blaming behavior coincides with their model of simple anger.

3 Before we conducted the experiment, the experimental design, the hypotheses and the planned analysis were pre-registered at AsPredicted and they can be found at the link http://aspredicted.org/blind.php?x=ij3fg3.
the strategy method, or at least it should be significantly attenuated compared to the direct method.

Subjects are assigned to one of four treatments, characterized by the method of play and the responder’s (player B) monetary payoff from the default allocation, which is either $m^1_b$ or $m^2_b$ with $m^2_b > m^1_b$. In every treatment, prior to choices in the experimental game, they face a prediction task through which we elicit beliefs of both proposers and responders about the behavior of subjects playing in the opponent’s role. Since frustration is rooted in unfulfilled expectations, our behavioral predictions about behavior in the experimental game across payoff treatments depend on the responder’s initial belief about the default allocation. Such belief is endogenous and may vary with the treatment. The effect of the payoff increase crucially depends on whether frustration actually increases with the responder’s payoff from the default allocation, which in turn depends on whether the subjective probability of a default allocation increases (or decreases only slightly).

We derive implications about behavior in the prediction task (beliefs) and in the experimental game (distribution of actions) from a model that assumes that proposers are self-interested, while responders have belief-dependent preferences given by a combination of anger and inequity aversion. The assumption that proposers are selfish is just a simplification. Indeed, BDS theory implies that first movers in a game cannot be angry. As for inequity aversion, it plays a secondary role in our model; furthermore, given the material payoffs in our treatments, inequity aversion matters more for responders than proposers. Our theoretical analysis takes into account that the anonymous interaction of experimental subjects necessarily features incomplete information about co-players’ relevant traits. In addition, the textbook equilibrium hypothesis that players have correct conjectures about the co-players’ decision rules is unjustified in this context. Therefore, we just assume that players perform two steps of elimination of non-best replies given plausible assumptions about beliefs. Specifically, we assume that the distributions of the responder’s initial beliefs in the two payoff treatments are such that responders in treatment $m^2_b$ do not expect that the default allocation is less likely than responders in treatment $m^1_b$. Moreover, we assume that the relation between the distributions of the proposer’s second-order beliefs in $m^1_b$ and $m^2_b$ reflects the relation between the distributions of the responder’s first-order beliefs in the two treatments. Finally, each player is rational—i.e., he best responds to his beliefs—and he is certain of the other’s rationality at the root of the game.

Given our experimental design, these assumptions on preferences and strategic thinking lead to testable predictions about beliefs and behavior. Our assumption about beliefs implies that B’s initially expected payoff is larger in payoff treatment.
than in \( m_{1b} \) (for fixed method of play). A’s certainty of B’s (subjective) rationality and the assumption about the distributions of the proposer’s second-order belief imply that—in the direct method—proposers in treatment \( m_{2b} \) believe that rejections are more likely than in treatment \( m_{1b} \). Predictions about beliefs have clear behavioral implications in the experimental game, due to the rationality assumption. In particular, with the direct method, we (like proposers) expect a higher conditional frequency of rejections in \( m_{2b} \) than in \( m_{1b} \). Moreover, we expect that greedy offers negatively correlate with the increase of the responder’s payoff from the default allocation (given beliefs). The anger model of BDS, independently of strategic considerations, also predicts (with the direct method) a positive correlation between rejections and responders’ initial beliefs in the default allocation.

Importantly, the behavioral predictions of our model hinge on whether the responder can experience frustration, which is not the case in the strategy method. Notice that, given our assumption on the responder’s preferences, our model admits a positive rejection rate in both methods of play and payoff treatments due to the inequity aversion component. However, the frequency of rejections is expected to be constant across payoff treatments in the strategy method since the frustration component is absent. By adopting a broader definition of frustration—different from the one adopted in BDS—and admitting that responders can to some extent image the feeling of frustration, we can in principle allow for a positive effect of the payoff increase on rejections in the strategy method. However, as argued above, imagining oneself being frustrated could not produce the same angry response as actually feeling frustrated. Thus, we expect that the increase of rejections due to the payoff increase in the strategy method is either absent or lower than in the direct method. The assumption that proposers are selfish is also relevant. If they have social preferences (e.g., inequity aversion), an increase in the default-allocation payoff of the responder (only) may make proposers more prone to make the greedy offer. Since we measure proposers’ beliefs, such counterbalancing effect may be revealed by the data.

In line with the theory of frustration and anger, in the direct method we find that the higher is the responder’s initial expectation of a default allocation, the less likely it is that he will accept a greedy offer. However, in contrast with our predictions, we find that increasing the responder’s payoff from the default allocation does not significantly increase the conditional frequency of rejections of the greedy offer nor the frequency of default allocations, independently of the method of play.

---

4Notice that this prediction relies on an important assumption, i.e., that the responder’s sensitivity to anger is constant across payoff treatments. This guarantees that the responder’s expected psychological utility from a greedy offer is larger in treatment \( m_{1b} \).
In line with our predictions, in the direct method subjects display a higher rejection rate than in the strategy method, but varying the method of play does not have a significant effect on the frequency of default allocations.

2 Related Literature

Besides reporting the closest literature on frustration and anger, we consider (1) the experimental literature focusing on the role of initial expectations in games that involve costly punishment actions, and (2) the experimental literature (both in economics and in psychology) that studies the role of emotions in Ultimatum-like games. We then explain how we link these two strands of the literature and how we innovate on them.

To the best of our knowledge, Persson (2018) is the only published article that tests a version of BDS theory of frustration and anger. In the context of a game where unexpectedly low payoffs are determined by the simultaneous uncoordinated choices of chance and of another player, he finds that, while unfulfilled expectations about material payoffs generate negative emotions (as in BDS theory), these emotions do not affect subjects’ punishing behavior. This result does not support some specifications of BDS theory, either simple anger, or the “could-have-been” version of anger from blaming behavior, both of which hinge on how exactly frustrated players blame their co-players. By testing BDS theory in the context of the UMG, we instead can focus on the core of BDS theory, without addressing specific versions of blaming.

Our paper relates to the experimental literature on second-party punishment. In particular, a strand of this literature focuses on the role of expectations as a crucial determinant of costly punishment actions in laboratory experiments. A series of papers investigating the Power-To-Take game (Bosman and van Winden 2002,

5There are also two recent unpublished studies by Dufwenberg et al. (2019a,b) that explore the role of communication—respectively promises in hold-up games and threats in deterrence games—when subjects are motivated by frustration and anger as in BDS theory.

6Another difference between our work and Persson (2018) is that, in his theoretical model, he assumes complete information about the punisher’s anger sensitivity and predicts behavior using equilibrium analysis, whereas we assume incomplete information and use two steps of elimination of non-best replies. However, complete-information equilibrium analysis plays a minor role in Persson’s derivation of behavioral predictions.

7Empirical studies show similar findings. In particular, it has been found that there is a correlation between unexpected losses suffered by local teams and the number of reports of domestic violence (Card and Dahl 2011) and violent crime (Munyo and Rossi 2013). According to BDS theory the unexpected negative event—i.e., the loss of the favorite football team—can be interpreted as an external source of frustration.

8In this game, the first-mover can appropriate part of the second mover’s income. In return, the second mover can retaliate destroying partially or fully his own resources.
Bosman, Sutter, and van Winden (2005, Reuben and van Winden 2008) finds that expectations on the take rate affect the decision to retaliate. In the experiment of Bohnet and Zeckhauser (2004), subjects playing in the role of responders of an Ultimatum Game are informed about the average amount offered in their session prior to their decision to accept or reject an offer. Receiving this information positively affects the probability of offer-specific rejections. Similar findings are found in Sanfey (2009) where receivers’ expectations are directly manipulated by the provision of information about amounts offered in previous sessions.

Results from neuroeconomics studies confirm the crucial role of expectations. By eliciting expectations, Chang and Sanfey (2013) find that players are more likely to reject offers that deviate from their expectations and that the magnitude of these deviations correlates with increased activity in brain areas associated with error-monitoring. Xiang, Lohrenz and Montague (2013) manipulate responders’ expectations varying the distribution of offers that they receive and find that rejections occur more often when subjects expect higher offers. Moreover, the magnitude of deviations correlate with areas related to negative emotions and also with areas related to the reward system.

Another strand of the literature on second-party punishment focuses on the role of subjects’ preferences. In economics, most of the experimental literature associates costly punishment behavior to inequity aversion and fairness considerations. This is due to the fact that the observed frequency of rejections usually increases as offers decrease in magnitude. Recently, a stream of literature has argued that motivations other than inequity-aversion may explain these findings. The experimental study on the Ultimatum Game by Xiao and Houser (2005) shows that the possibility of expressing negative emotions (anger and disapproval) directly to proposers in reaction to greedy proposals decreases the frequency of rejections. In a similar experiment, van Kleef et al. (2004) show that proposers become more generous towards responders who express anger-like emotions. A recent study by van Leeuwen et al. (2017) confirms this result and provides experimental support to the hypothesis that anger serves as a credible commitment device (Frank 1987, 1988). First, responders who reject low offers display anger facial expressions (measured by a face-reader software). Second, external observers are able to detect better than a random device

\[9\] Notice however that the authors do not actually interpret rejections as driven by unfulfilled expectations, but as reactions to the violation of the norm of equity (induced by the provision of information about the average offer in the same session).

\[10\] A similar effect of the strategic use of anger is found by Gneezy and Imas (2014) in an experiment involving two-player zero-sum games. Interestingly, subjects seem to anticipate the emotional reaction of their co-players and exploit the strategic opportunity to anger their opponents in scenarios where this can be advantageous.
subjects who then rejected greedy allocations, on the basis of responders’ pictures only, taken prior to receiving any information about the game.

In addition, the intriguing experiments of Grimm and Mengel (2011) and Oechssler et al. (2008) show that time may mitigate negative emotions arising with greedy offers: delaying the responder’s decision increases the acceptance rate for low offers. Gneezy and Imas (2014) similarly find that waiting reduces anger. On the other hand, Sutter et al. (2003) show that forcing the responder to answer within a short time frame improves rejection rates. Also, the rejection rate in the UG is higher when the allocation is decided by a person with respect to the case where an algorithm splits the money randomly between the two players (Blount 1995). All these studies suggest not only that negative emotions like anger play a role in Ultimatum-like games, but also that models of social preferences developed within the standard consequentialist framework of decision theory cannot fully account for rejections, as the time of response should not have any effect on behavior. Instead, BDS theory for multi-period models accounts for the mitigating effect of delay (see Battigalli et al. 2019a, Sections 2.3 and 3.3).

A strand of the psychological literature also provides empirical support for anger as the underlying mechanism for costly punishment. Pillutla and Murnighan (1996), relying on self-reports of emotions, find that low offers are often rejected and interpret rejections as angry reactions to perceived unfairness. Few subjects report anger alone, with no association to perceived unfairness. Moreover, rejections are better predicted by reported anger than perceived unfairness. Seip et al. (2014) find that, both in a public good game and in a trust game, the lower the level of cooperation by co-players, the more anger is evoked and the harsher the punishment. In addition, in the trust game, increased anger (manipulated through an unrelated task) produces more severe punishments. Notice that, similarly to BDS theory, Seip et al. (2014) interpretation is also derived by the appraisal theory and anger is interpreted as the action tendency of a negative emotion. Yet, while in Seip et al. (2014) anger derives from the appraisal of co-player’s unfairness, in BDS theory anger derives from the appraisal of an unexpected event that blocks one’s goals.

Furthermore, several neuroeconomics experiments employ the use of functional magnetic resonance imaging (fMRI) to study the neural activations of cognitive emotional processes in the Ultimatum Game. These studies reinforce the idea that rejections of lower offers are linked to higher activity of brain regions related to emotional decision making (Gospic et al. 2011, Gilam et al. 2018) and in particular to negative emotions (Sanfey et al. 2003, Gabay et al. 2014). Specifically, Gilam

---

11Similar studies (van Wout et al. 2006, Civai et al. 2010) analyze skin conductance, a measure of emotional arousal, of responders who are asked to accept or reject a series of offers. Results
et al. (2018) study the psychometric properties of both a standard and an anger infused Ultimatum Game—where interpersonal provocations are allowed—and assess their validity as anger induction paradigms. In both games they find a negative correlation between the probability of acceptance (which is positively correlated with offer fairness) and reported anger, modulated by brain activity in the ventro-medial prefrontal cortex, which is key in emotion regulation.\(^{12}\)

All this evidence taken together suggests that there is a relation between expectations and rejections and that negative emotions like anger play a role in Ultimatum-like games. Yet, none of the papers cited above provides a theoretical interpretation that can jointly account for the role of expectations and anger in these social dilemmas, as we do in this paper.\(^{13}\) Building on previous evidence, we interpret rejections through the lenses of BDS theory that links expectations, anger-like emotions and costly punishment actions. As we assume that the responder’s decision utility is belief-dependent, we depart from the standard consequentialist framework of decision theory that instead characterizes earlier interpretations for rejections in the Ultimatum Game.\(^{14}\) Moreover, the assumption of belief-dependence informed our experimental design which is rigorously motivated by a theoretical model. Importantly, previous interpretations could not explain differences in the responders’ behavior across different methods of play—i.e., the direct response and the strategy method—that have found support in experimental data and that can actually be explained by belief-dependence.

Finally, we deviate from most of the related literature in the way we model strategic thinking. Our behavioral predictions across treatments are neither derived from equilibrium analysis, nor merely from correlations between behavior and elicited beliefs. Following Battigalli et al. (2013) we instead derive comparative predictions from two steps of elimination of non-best replies, given some assumptions about agents’ beliefs. Unlike Battigalli et al. (2013), our assumptions compare distributions of beliefs in different treatments.

\(^{12}\)Subjects’ rejections in the anger infused version of the Ultimatum Game display a stronger relationship with reactive aggression (as measured by a psychometric task) and with trait anger—i.e., the subjects’ individual habitual tendency to experience anger—measured through the STAXII (Spielberger 1999). Thus, the anger infused version outperforms the basic version of the game as a more appropriate mechanism to induce interpersonal anger and its aggressive expression for basic and clinical research settings.

\(^{13}\)As explained above, Persson (2018) tests BDS theory, but in a different strategic situation. Furthermore, he assumes equilibrium and complete information.

\(^{14}\)A different departure from the consequentialist interpretation of rejections in the Ultimatum game is the reciprocity motivation (Charness and Rabin 2002 and Falk et al. 2003). In Section 7, we will discuss the implications of reciprocity in our strategic setting and whether reciprocity can explain the evidence that we find.
3 Theory of Frustration and Anger in the UMG

We study the implications of the theory of frustration and anger in the strategic context of the UMG (Binmore et al. 1995), a simple binary-choice version of the Ultimatum Game (Güth et al. 1982). The game form with material payoffs is represented in Figure 1, where \( h > m_i > \ell > 0, \ i = a, b \). In this highly stylized social dilemma the first-mover can either propose a default allocation \((d)\), whereby both players receive a similar amount of money, or a “greedy” allocation \((g)\). While the default allocation is automatically accepted, the second-mover can either accept or reject the greedy offer.

![Figure 1: The Ultimatum Minigame](image)

To test experimentally BDS theory we consider two manipulations. Firstly, we change the payoff of the responder from the default allocation \((m_b)\). Specifically, we will consider two alternative payoff structures, one with \( m_a = m_b \) and the other with \( m_b > m_a \). The purpose is that of increasing the initially expected payoff of the responder without varying the proposer’s payoff from the default allocation. Secondly, we vary the method of play (direct response method versus strategy method) for reasons that will be clear in the analysis that follows. We let index P denote the method of play, with \( P = D \) if UMG is played with the direct method and \( P = S \) if it is played with the strategy method.

For a purely material payoff maximizer in the role of player B who has received a greedy offer, the only rational choice is to accept it and get a positive payoff \((\ell > 0)\). However, experimental evidence shows that responders do not always accept unequal allocations even when, by rejecting them, they forego a material gain (e.g., Güth et al. 1982, Hoffman et al. 1996, Slonim and Roth 1998, Cameron 1999, Carpenter et al. 2005, Andersen et al. 2011).

In what follows we explain how BDS theory can explain rejections in UMG relying

\[ \text{In the classic version, the two players are given an amount of money to split, thus it has to be the case that } m_a + m_b = h + \ell. \]
on the psychological preferences of the responder. The intuition is the following. Player B may feel frustrated after observing the greedy offer if the maximal payoff that he can achieve conditional on this offer is lower than what he expected to obtain at the beginning of the game. Frustration triggers anger that, in turn, depending on player B’s sensitivity to this emotion, can lead to a rejection. The higher player B’s initial expectations at the root of the game, the more frustrated he is after observing the greedy offer and thus the more willing to reject it.

Models of social preferences are usually invoked to explain rejections in the Ultimatum Game. The most prominent of them is the inequity aversion model by Fehr and Schmidt (1999). Inequity averse preferences predict a positive rejection rate in UMG: an inequity averse player B is willing to reject the greedy offer if he prefers the equal zero-payoff allocation over the unequal one \((h, \ell)\). Notice that in the inequity aversion model—as in other models of distributional preferences—player B’s decision utility is purely outcome-based.\(^\text{16}\) Hence, only the comparison between the two outcomes available after the greedy offer matters. Instead, according to BDS theory player B’s payoff from the default allocation can influence his decision as it determines his initially expected payoff and thus, the extent of his frustration after the greedy offer (see Battigalli et al. 2019b, Example 3).

Since the theory that drives our experiment hinges on the psychological preferences of player B, who is the only player that can be affected by frustration and anger,\(^\text{17}\) here we focus on player B’s decision conditional on the greedy offer and derive the behavior strategy associated to his psychological utility. In order to relate the behavioral implications of the theory of frustration and anger with the predictions of the most representative model of social preferences, we embed the psychological component of frustration and anger into a model that encompasses also the inequity aversion motive. Thus, we assume that player B’s decision utility function is affected both by anger (Battigalli et al. 2019b) and inequity aversion (Fehr and Schmidt 1999).\(^\text{18}\)

We denote with \(\beta\) the initial first-order belief of player B about player A choosing

\(^{16}\)For a discussion of the difference between decision and experienced utility please see Battigalli et al. (2019a), Section 6.

\(^{17}\)According to BDS theory, only the second-mover can be frustrated. Indeed, unlike other emotion-based preferences, BDS definition of anger is based on an action tendency that is triggered by the observation of an event that was initially unexpected (Lerner and Keltner 2000, 2001). In fact, the proposer cannot be motivated by this emotion as, being the first to move, there is no event that precedes his decision whose occurrence can cause his frustration.

\(^{18}\)This model of inequity-aversion proposes a utility function for player \(i\) of the general form:

\[u_i(x) = x_i - \delta_i \max\{0, x_j - x_i\} - \eta_i \max\{0, x_i - x_j\}\]

with the sensitivity parameters such that \(\delta_i \geq \eta_i\) and \(\eta_i \in [0, 1)\). However, in our context it is never the case that \(x_i > x_j\); therefore, the last term is missing in our model.
the default allocation, i.e., $\beta = P_b(d)$, and with $\gamma$ B’s initial first-order belief about himself accepting the greedy offer, i.e., B’s planned probability of acceptance $\gamma = P_b(y|g)$. B’s first-order beliefs $\beta$ and $\gamma$ affect his decision utility through frustration.

According to BDS theory, frustration is anchored in the appraisal of a negative event, which is, in this context, the greedy offer. Thus, player B can be frustrated and blame player A only after actually observing the greedy offer. Notice that, when UMG is played with the direct method ($P = D$), player B observes the greedy offer before taking his decision. Instead, when UMG is played with the strategy method ($P = S$), player B commits to a decision rule that selects the reply conditionally on the hypothesized greedy offer, hence he cannot be frustrated. Therefore, while in the strategy method frustration is equal to zero, in the direct method it is the gap between player B’s expected payoff at the root of the game—determined by $\beta$ and $\gamma$—and the maximal payoff that he can achieve after the greedy offer, i.e., $\ell$:

$$F_{P,m_b} (\beta, \gamma) = \begin{cases} \max \{0, \beta m_b + (1 - \beta) \gamma \ell - \ell \} & \text{if } P = D \\ 0 & \text{if } P = S. \end{cases}$$ (1)

In the direct method, player B is more frustrated (i) the more he expects a default allocation, i.e., the higher $\beta$, and (ii) the more he plans to accept the greedy offer at the beginning of the game, i.e., the higher $\gamma$.

Next, we specify B’s belief-dependent psychological utility of replying with action $a_b$ to the greedy offer $g$ as

$$u_{P,m_b}^b (g; a_b; \beta, \gamma) = \pi_b (g, a_b) - \pi_a (g, a_b) \theta F_{P,m_b} (\beta, \gamma) - \delta \max \{0, \pi_a (g, a_b) - \pi_b (g, a_b)\},$$ (2)

where $\pi_i (g, a_b)$, $i = a, b$, is the monetary payoff of player $i$ after the greedy offer and B’s choice $a_b$ ($a_b \in \{y, n\}$), $\theta$ and $\delta$ are B’s personal traits which measure his sensitivity to anger and his inequity-aversion, that is the distributive component of his preferences.

Hence, considering the two methods of play separately and replacing $F_{P,m_b} (\beta, \gamma)$ with the corresponding expressions from equation (1), we obtain:

$$u_{P,m_b}^b (g, a_b; \beta, \gamma) = \begin{cases} \ell - \delta (h - \ell) & \text{if } a_b = y \\ 0 & \text{if } a_b = n. \end{cases}$$

and

$$u_{D,m_b}^b (g, a_b; \beta, \gamma) = \begin{cases} \ell - h \theta \max \{0, \beta m_b + (1 - \beta) \gamma \ell - \ell \} - \delta (h - \ell) & \text{if } a_b = y \\ 0 & \text{if } a_b = n. \end{cases}$$
In words, since we assume that player B cannot be frustrated in the strategy method from “hypothesizing” the greedy offer, the frustration channel is shut down and the resulting (inequity averse) preferences are consequentialist. Therefore, B’s utility from accepting the greedy offer does not depend on his first-order beliefs \( \beta \) and \( \gamma \), nor on his payoff from the default allocation. Hence, in the strategy method player B accepts the greedy offer if and only if

\[
u_b^{S,m_b}(g, y; \beta, \gamma) = \ell - \delta (h - \ell) \geq 0 = u_b^{S,m_b}(g, n; \beta, \gamma),
\]

that is,

\[
\delta \leq \delta^S := \frac{\ell}{h - \ell}.
\]

Player B’s initial expectations and sensitivity to anger kick in when the game is played with the direct method. In this case, player B accepts the greedy offer if and only if

\[
u_b^{D,m_b}(g, y; \beta, \gamma) = \ell - h\theta F^{D,m_b}(\beta, \gamma) - \delta (h - \ell) \geq 0 = u_b^{D,m_b}(g, n; \beta, \gamma).
\]

To provide a unified representation of the best reply correspondence, we express acceptance as a condition on \( \delta \) also for the direct method, but in this case the threshold depends on B’s belief \( \beta \), his planned probability of acceptance \( \gamma \), and his sensitivity to anger \( \theta \):

\[
\delta \leq \delta^{D,m_b}(\beta, \gamma, \theta) := \frac{\ell - h\theta \max \{0, \beta m_b + (1 - \beta) \gamma \ell - \ell\}}{h - \ell},
\]

(3)

Note that, differently from the strategy method, the threshold depends on B’s payoff from the default allocation, which influences his initially expected payoff, and thus his level of frustration after the greedy offer.

The following proposition summarizes:

**Proposition 1.** Given the method of play \( P \) and B’s payoff from the default allocation \( m_b \), player B accepts the greedy offer if and only if his degree of inequity aversion is low enough, i.e., if and only if

\[
\delta \leq \delta^{P,m_b}(\beta, \gamma, \theta) := \frac{\ell - h\theta \max \{0, \beta m_b + (1 - \beta) \gamma \ell - \ell\}}{h - \ell} = \begin{cases} \frac{\ell - h\theta \max \{0, \beta m_b + (1 - \beta) \gamma \ell - \ell\}}{h - \ell} & \text{if } P = D \\ \frac{\ell}{h - \ell} & \text{if } P = S. \end{cases}
\]

The proposition shows that the acceptance condition in the strategy method is easier to meet than in the direct method. Furthermore, changes in the default allocation payoff \( m_b \) can affect B’s behavior only when UMG is played with the direct method. When \( m_b \) increases, other things being equal, B’s expected payoff
also increases and so does his frustration, making it more difficult to meet the acceptance condition \[3\].

Figure 2: B’s intrapersonal equilibrium correspondence as a function of \(\beta\) and \(\delta\), with parameters \(\theta = 0.015, h = 14, m = 8, \ell = 2\).

When \(P = D\) and anger bites, B’s best reply to the greedy offer also depends on his plan of accepting it. In particular, for some payoff conditions and personality features, a plan to accept \((\gamma = 1)\) may induce high frustration and rejection, and a plan to reject \((\gamma = 0)\) may induce low frustration and acceptance. In both cases, the plan is not incentive compatible and a fully rational responder would not have such plan. In order to derive qualitative predictions on B’s behavior, we assume that he has “rational expectations” about his own behavior, that is, he correctly predicts that he is going to choose a best reply. In other words, B’s behavior strategy \(\gamma\) has to satisfy the one-shot deviation property and thus yields an intrapersonal equilibrium for B.\[19\] For given payoff conditions, there are three cases depending on B’s personal features (see Figure 2):

- for low values of \(\beta, \theta,\) and \(\delta\), anger and inequity aversion are low and the plan to accept is incentive compatible \((\gamma = 1)\);

- for intermediate values of \(\beta, \theta,\) and \(\delta\), deterministic plans are self-defeating, i.e., a plan to accept induces rejection and a plan to reject induces acceptance via the effect of \(\gamma\) on the initially expected payoff and frustration; thus the only incentive compatible plan is a probability of acceptance \(\gamma \in (0, 1)\) that makes B indifferent;

\[19\] The intrapersonal equilibrium approach is standard in the analysis of sophisticated agents with dynamically inconsistent preferences. See, for example, the models of Caplin and Leahy (2001), Koszegi and Rabin (2009) and the methodological contributions by Battigalli and Dufwenberg (2009) and Battigalli et al. (2019a). Note that this is only an intrapersonal equilibrium because belief \(\beta\) is not assumed to be correct.
for high values of $\beta$, $\theta$, and $\delta$, anger and inequity aversion are high and the plan to reject is incentive compatible ($\gamma = 0$).

We prove in the Appendix that, neglecting non-generic cases that make $\gamma$ indeterminate, B’s intrapersonal equilibrium correspondence is as follows (see left panel of Figure [2]):

$$\Gamma^D,mb(\beta, \theta, \delta) = \begin{cases} 
1 & \text{if } \delta < \delta^D,mb(\beta, 1, \theta) \\
\frac{\ell - h(\beta m_b - \ell) - \delta (h - \ell)}{h(1 - \beta)\ell} & \text{if } \delta \in \left[\delta^D,mb(\beta, 1, \theta), \delta^D,mb(\beta, 0, \theta)\right] \\
0 & \text{if } \delta > \delta^D,mb(\beta, 0, \theta),
\end{cases}$$

where $\delta^D,mb(\beta, 1, \theta) \leq \delta^D,mb(\beta, 0, \theta)$ because frustration—hence the incentive to reject—is decreasing in $\gamma$.

The following lemma establishes that the incentive-compatible probability of accepting the greedy offer decreases with B’s expectation of the default allocation.

**Lemma 1.** $\Gamma^D,mb(\beta, \theta, \delta)$ is decreasing in $\beta$.

This result characterizes the dependence of player B’s behavior on his initial expectations. The higher his first-order belief about the default allocation (given $m_b$), the larger it is his initially expected payoff, and thus his frustration after the greedy offer. A higher frustration triggers a stronger angry reaction, increasing B’s willingness to reject the greedy offer.

### 4 The Experiment

In this section we describe our experiment. First, we describe the treatments in details. Next, we describe the experimental procedures.

#### 4.1 Experimental Design

We conduct a two-fold manipulation, informed by the analysis of player B’s behavior strategy reported in the previous section. The experimental design was pre-registered on the AsPredicted platform, together with the experimental hypotheses and the analysis plan.

**Payoff manipulation.** With the purpose of increasing player B’s frustration, we manipulate player B’s material payoff from the default allocation ($m_b$), while

keeping all the other payoffs constant. Under some assumptions on player B’s initial beliefs, such payoff manipulation results in an increase of B’s initially expected payoff, and thus of his frustration in case of a greedy offer. Indeed, increasing $m_b$ increases (ceteris paribus) the gap between his initially expected payoff and $\ell$, the maximal payoff that he can achieve upon receiving a greedy offer. Notice that when varying $m_b$, player B’s initial beliefs may also change—i.e., beliefs are endogenous to the treatment. Therefore, player B’s initially expected payoff actually increases provided that he believes that the default allocation is not less likely when his payoff from this offer is larger.

As player B’s best reply does not vary with player A’s payoff from the default allocation ($m_a$) we keep this payoff constant. At the same time, we do not alter A’s and B’s payoff from accepting the greedy offer ($h$ and $\ell$) in order to keep constant the effect that inequity aversion may have on B’s behavior across payoff treatments. Thus, while in a payoff treatment B-subjects obtain $m_b^1$ from the default allocation, with $m_b^1 = m_a$, in the other payoff treatment they obtain $m_b^2 = m_b^1 + \varepsilon$.

**Method of play manipulation.** With the purpose of switching on and off player B’s experience of frustration, we manipulate the method of play, i.e., whether UMG is played with the direct response method ($P = D$) or the strategy method ($P = S$).

The effect of the method of play has been debated in experimental economics for long time (Brandts and Charness 2011). Experimental studies on the Ultimatum Game have shown mixed evidence of differences across methods of play due to the high potential of the direct response method in triggering emotional responses. According to the standard theory of decision under uncertainty, the method of play should not alter player B’s behavior and thus it should not influence the effect of the payoff increase, as, in this framework, B’s initial expectations are irrelevant for his decision. Indeed, for an expected utility maximizer strategies that maximize subjective expected utility ex ante are also optimal conditional on any information that the decision maker deems possible. Even in the case of belief-dependent preferences, as long as psychological utility is independent of the agent’s own plan, the method

---

21 Falk et al. (2003) investigates various versions of the UMG; in particular, they compare the rejection rate in a traditional UMG, where the default allocation is (5,5) and the offer is (8,2) with an UMG where the default allocation is (2,8) and the offer is (8,2). In the latter treatment, where the default option is very unfair to the proposer, they find that the rejection rate drops from 44% to 27% and the choice of the offer (8,2) grows from 31% to 73%. We instead keep the proposer’s payoff constant in the default allocation and slightly increase the responder’s payoff in order to minimize the effect that varying the distribution of payoffs in the default allocation may have both on offers and rejections. Indeed, our objective is to isolate anger as the driving force of rejections and—if anticipated—of making less unfair offer.

22 Gith et al. (2001) find that varying the payoff distribution of the outside option affects proposers’ behavior only in the direct method, while (probably due to the few observations) they do not find any effect of the method of play on rejections.
of play is unimportant, because the traditional dynamic consistency property of expected utility theory still holds (see Battigalli et al. 2019a, Section 6 and Remark 1 of Section 7).

According to BDS theory instead the method of play may affect behavior. Indeed, under a psychological perspective, it makes a difference whether player B’s choice is the action tendency of an emotion that is triggered by the appraisal of an event whose occurrence is actually observed, as in the direct response method, or only imagined, as in the strategy method. As discussed in Section 3, when $P = S$ player B’s behavior is invariant to $m_b$ since he cannot experience frustration—i.e., the anger component does not bite—and player B’s decision only depends on his degree of inequity aversion.

Obviously, we can in principle admit that player B is to some extent able to anticipate the negative emotion that he would experience in case he received a greedy offer. Yet, it is rather implausible that this emotion can be so strong to trigger an action that makes him forego a private gain for the goal of retaliating against player A’s hypothetical move. Thus, in the strategy method player B’s frustration is expected to be at least attenuated with respect to the direct method and so it is the effect of the payoff increase.

Table 1 summarizes our $2 \times 2$ design, reporting our four treatments and the corresponding labels.

<table>
<thead>
<tr>
<th>Payoff treatment $m_{b1}$ $m_{b2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct response method D1, D2</td>
</tr>
<tr>
<td>Method of play</td>
</tr>
<tr>
<td>Strategy method                  S1, S2</td>
</tr>
</tbody>
</table>

We implement a between-subjects design, so that subjects play only one of the treatments D1, D2, S1, and S2. Figure 3 illustrates the UMG game form for the two payoff treatments, that differs only for player B’s payoff from the default allocation, which is 8 in payoff treatment $m_{b1}$ and 11 in payoff treatment $m_{b2}$.

Before subjects play the UMG we elicit their first-order beliefs with a procedure that we will explain in detail below. In our context, eliciting beliefs is particularly important for two reasons. The first is that by doing so we can verify whether B-subjects initially expect more the default allocation in treatment D2 than in treatment D1. Secondly, it enhances our understanding of subjects’ strategic reasoning.
4.2 Experimental procedures

We recruited 352 participants from the subjects pool of Bocconi University students with the software SONA. The experiment was programmed with the software oTree (Chen et al. 2016) and run in the BELSS laboratory of Bocconi University. Average pay was 8.70 EUR (including a show up fee of 4 euro) and the experiment lasted on average 45 minutes. We ran 16 sessions, with 11 pairs of subjects per session, formed randomly and anonymously: to have a comparable number of observations for the responders, we had 3 sessions for each payoff treatment in strategy (33 observations per role) and 5 sessions for each payoff treatment in direct, resulting in 55 observations per payoff treatment for A and 72 for B, 35 in D1 and 37 in D2, respectively). Instructions were read aloud and we made sure subjects understood the rules by asking control questions.

We elicited first-order beliefs of both A and B-subjects before choices with a very simple procedure. As in every session there were 11 pairs of subjects, we asked B-subjects to guess how many A-subjects of the other pairs in the session would choose the default allocation. Similarly, in the strategy method, we asked A-subjects to guess how many B-subjects of the other pairs would choose to accept the greedy offer. Instead, in the direct method, we asked A-subjects to guess which percentage of B-subjects, among those who received a greedy offer, would decide to accept it. All subjects were paid 10 ECUs in case of a correct guess and 2 ECUs in case of a wrong guess.

---

23 The subject pool includes undergraduate students from different disciplines (Economics, Business, Law and Political Science).

24 Whether it is better to elicit beliefs before or after choices has been highly debated in the experimental economic literature. The evidence found by Nyarko and Shotter (2002), Costa-Gomes and Weizsacker (2008) and Ivanov (2011), showing that the impact of ex ante belief elicitation on subjects’ behavior is negligible, supports our design choice. Besides all technical considerations that might be put forward in favor of the choice of eliciting beliefs ex ante, in our context it is compelling to do so, as BDS theory relates player B’s choice to his initial expectations—i.e., his expectations at the root of the game.
guess. Moreover, with the purpose of avoiding hedging, subjects were either paid for the beliefs elicitation task or for the UMG play (with the same probability). Notice that asking subjects to make a guess about the behavior of subjects’ playing in the opponent’s role in the other pairs has the purpose of avoiding additional (and confounding) emotional responses of B-subjects in the direct method. In case they had to make guesses also on their co-player’s behavior, after observing the greedy offer, they might be disappointed for realizing that they might have made a wrong guess in the beliefs task.

After the experiment, we asked B-subjects to fill the State-Trait Anger Expression Inventory-2 questionnaire, namely STAXI-II (Spielberger 1999), that delivers both a measure of sensitivity to anger as a stable trait and a measure of the anger state. While B-subjects were filling the STAXI-II, A-subjects were asked to fill the Aquino questionnaire about morality (Aquino and Reed 2002) and the GASP test of guilt aversion (Coehn et al. 2011). Notice that while the STAXI-II is important as it helps us have a measure of B-subjects’ sensitivity to anger and understand the motivation behind their behavior, the questionnaires administered to A-subjects are less relevant and their main purpose was to make the duration of the experiment symmetric across A and B-subjects. Yet, both questionnaires measure personal traits that potentially affect A-subjects’ behavior and any correlation between this measures and A-subjects’ choices may contrast the selfishness assumption.

In the next section we derive our behavioral predictions, given player B’s behavior strategy analyzed in Section 3 and the experimental design described in this section.

5 Behavioral Predictions

While in Section 3 we have studied the determinants of player B’s decision, our goal in this section is to obtain empirical predictions about how the distribution of actions in UMG changes across treatments—i.e., how it changes depending on player B’s payoff from the default allocation and the method of play. We first provide qualitative predictions about B and A’s behavior across payoff treatments in the direct method. Indeed, according to BDS theory, only in this condition player B’s preferences are belief-dependent and can be influenced by variations in his payoff from the default allocation, as extensively discussed in the previous sections. Next,

\[\text{Obviously, subjects in the same pair were paid for the same task, in order to provide them with appropriate incentives within the UMG play.}\]

\[\text{26 The Aquino Moral Identity Scale measures the degree to which people's self-concepts center on moral traits—i.e., how much people consider their morality to be central to their identities.}\]

\[\text{27 The GASP test measures individual differences in the propensity to experience guilt and shame across a range of personal transgressions.}\]
we analyze B and A’s behavior across methods of play (direct and strategy), holding constant the payoff treatment.

In traditional economic theory, UMG admits only one subgame perfect equilibrium in which a selfish player A makes the greedy offer and a selfish B accepts. Here we assume instead that subjects may be affected by frustration and anger, which can happen only when they play in role B. Furthermore, we assume that B-subjects—unlike A-subjects—may be inequity averse. With this, we proceed with an analysis based on an iterated elimination of non-best replies under plausible assumptions about beliefs.

Most economic models rely on the assumption that players’ beliefs are coordinated on an equilibrium, hence that they are correct. Such equilibrium assumption is sometimes a useful theoretical shortcut, but it is justified only under special circumstances, that is, either when agents play the same game recurrently and it can be shown that learning eventually leads to correct beliefs or when strategic reasoning based on whatever is common knowledge about the game yields a unique prediction. Neither of these conditions hold in our experiment: Agents have no way to learn; furthermore, agents’ non-selfish or belief-dependent preferences are not common knowledge, which implies that strategic reasoning cannot yield unique predictions (see the discussion and analysis in Battigalli et al. 2019, and Attanasi et al. 2013). However, giving up the equilibrium assumption is far from giving up on predictive power, because even if players fail to coordinate on an equilibrium, they may be rational and confident in others’ rationality. This minimalistic approach allows to make qualitative predictions across treatments without relying on the equilibrium assumption.

We focus on a setting that mimics the random matching structure typical in laboratory experiments, as if experimental subjects are the players of a population game implemented by the experimenter in which subjects are randomly matched with one another. We analyze the game as an interaction between two individuals, player A and player B, that are drawn at random from a large population of subjects. As anticipated in Section 3, the relevant personal features of the individual playing in role B are \((\beta, \theta, \delta)\), i.e., his first-order beliefs and his personal traits. Instead, the relevant personal features of the individual playing in role A coincide with his

---

28 Thus, we assume that inequity aversion is triggered by playing in role B. Since the inequity aversion considered in this paper is the self-serving kind (only unfavourable inequality matters), we find this assumption a reasonable simplification.

29Subjective Bayes-Nash equilibrium is only a partial exception: beliefs about coplayers’ (endogenous) decision functions are assumed to be correct, but beliefs about their (exogenous) types are subjective, hence potentially incorrect. See Attanasi et al. (2016) for an analysis and discussion of subjective Bayesian equilibrium in a psychological game.

30See, e.g., Battigalli et al. (1992), and Fudenberg and Levine (1998).
personal beliefs. Player A’s relevant beliefs are his initial first-order beliefs about B accepting the greedy offer and his second-order beliefs about B’s initial expectation of a default allocation. Let \( \alpha = \mathbb{P}_a(y|g) \) denote the first-order belief of A. Denote by \( \phi \) the marginal second-order belief of A about B’s personal features, i.e., a joint belief on B’s first-order belief \( \beta \) and personal traits \((\theta, \delta)\) given by a subjective probability measure on \([0, 1] \times \mathbb{R}_+^2\)\(^{31}\). We will show how \( \alpha \) is derived from \( \phi \).

For each treatment, we posit a distribution of the relevant personal features of subjects, \((\beta, \theta, \delta)\) for B, and \((\alpha, \phi)\) for A. Of course, the marginal distribution of personal traits \((\theta, \delta)\) is exogenous—i.e., independent of the treatment—while beliefs are endogenous, as they are affected by strategic thinking. Since beliefs are treatment specific, also the thresholds for the personal trait \( \delta \) derived in the analysis of B’s behavior strategy (Section 3) vary with the treatment.

Given the random draw, the personal features of the individual playing in role A or B are vectors of random variables, which we denote with \textbf{boldface} letters. Thus, \((\beta_{Pk}, \theta, \delta)\) is the random vector of personal features of B, where \(\beta_{Pk}\) represents B’s beliefs in treatment Pk, with \(k \in \{1, 2\}\) indicating the payoff condition \(m^k_b\). The random vector of A’s personal features is \((\alpha_{Pk}, \phi_{Pk})\). Moreover, let \(\gamma_{Pk}\) denote B’s planned probability of acceptance of the greedy offer in treatment Pk.

Our behavioral predictions follow from some plausible assumptions about players’ beliefs. As in the strategy method preferences are not belief-dependent, assumptions about beliefs matter only for the direct method. In particular, we make across-treatment (D1 vs D2) comparative assumptions about the distributions of beliefs, which are meant to hold \textit{almost everywhere} given the distributions of personal features. We emphasize that we do not ascribe such comparative assumptions to the agents, i.e., we do not interpret them as coming from across-treatment reasoning, as such interpretation is prevented by our between-subjects setting.

We now explain in detail our assumptions about B-subjects. We assume that the distributions of B’s initial belief in the two payoff treatments satisfy \(\beta^{D2} \geq \beta^{D1}\), i.e., B-subjects in D2 tend to assign higher probability to the default allocation than B-subjects in D1. This assumption can be justified by an intuitive reasoning. The probability of proposing the default allocation reflects both the opportunity cost of choosing a safe amount \(m_a\) with respect to a larger one \(h\) given the concern that B may reject offer \(g\). While the first aspect is kept constant across payoff treatments, the second may be affected by the payoff increase as the gap between B’s monetary payoff in the two allocations is larger in payoff treatment D2. As we are taking in

\(^{31}\)Measure \(\phi \in \Delta([0, 1] \times \mathbb{R}_+^2)\) is a \textit{marginal} second-order belief because A is assumed to hold a joint belief about B’s \textit{action} and personal features \((\beta, \theta, \delta)\), which is the whole second-order belief of A.
consideration B’s first-order beliefs, it is plausible to assume that B-subjects in D2 are more prone to think that A is worried about a rejection relative to B-subjects in D1.

The first step of elimination of non-best replies is based on the following assumption.

**Assumption 1.** (i) Every player is subjectively rational, i.e., he plans rationally given his beliefs about the other, and implements his plan; (ii) B’s beliefs distributions in D1 and D2 satisfy $\beta^{D2} \geq \beta^{D1}$.

Notice that (i) refers to an elaboration of the standard notion of rationality with subjective expected utility maximization. Player B’s subjective utility coincides with his psychological utility given the greedy offer defined in condition 2 of Section 3. As we explained, such psychological utility depends on B’s planned probability of acceptance $\gamma$. Thus, we require that B plans in advance to take his psychological utility maximizing action with probability 1 whenever it is unique, and that he actually chooses with the planned probability when indifferent. This fixed-point condition is an “intrapersonal rational-expectations equilibrium” for B.

Assumption 1 yields comparative behavioral implications. Recall that $\Gamma^{Dk}(\beta, \theta, \delta)$ denotes the probability of accepting the greedy offer that satisfies the intrapersonal equilibrium condition given personal features $(\beta, \theta, \delta)$ in treatment Dk. Also, let $\mu^{Dk} \in \Delta ([0, 1] \times \mathbb{R}^2_+)$ denote the distribution of B’s personal features in treatment Dk, which we assume to be atomless. Notice that due to the rational expectations assumption—i.e., B’s beliefs about his own behavior are correct—the objective probability that B accepts the greedy offer coincides with the subjective probability that B attaches to this event. With this, the frequency of acceptance of the greedy offer is

$$P(a_b^{Dk} = y) = E[\gamma^{Dk}] = \int \Gamma^{Dk} d\mu^{Dk}.$$ 

The next proposition states the central behavioral prediction of our analysis.

**Proposition 2.** The frequency of acceptance of the greedy offer in treatment D1 is higher than in treatment D2, i.e.,

$$P(a_b^{D1} = y) \geq P(a_b^{D2} = y).$$

As explained above, this result is obtained under the assumption that B correctly predicts his own behavior, which allows us to pin down B’s behavior strategy as a function of his personal features $\beta, \theta$ and $\delta$. Yet, given the specification of monetary payoffs in our experiment, we can obtain the same prediction with an alternative assumption.
Remark 1. If $m_2^2 - m_1^1 > \ell$, i.e., the payoff increase is larger than player B’s payoff from the greedy allocation, and $\beta^{D_2} \geq 1/2 \geq \beta^{D_1}$ then the frequency of acceptance of the greedy offer in treatment $D_1$ is higher than in treatment $D_2$.

The intuition is the following. The larger payoff from the default allocation does not necessarily imply that B-subjects in $D_2$ are more frustrated than B-subjects in $D_1$, as in $D_2$ they may have a lower initial expectation to accept the greedy offer, that decreases the initially expected payoff. The two conditions of Remark 1 guarantee that even in the worst case scenario where the expectation of accepting the greedy offer in $D_1$ ($D_2$) is highest (lowest) and thus the expected payoff in $D_1$ ($D_2$) is maximal (minimal)—i.e., $\gamma^{D_1} = 1$ and $\gamma^{D_2} = 0$—the relation between the expected payoffs in the two treatments is satisfied for every realization of $\gamma^{D_1}$ and $\gamma^{D_2}$. Since the decision of acceptance in each treatment depends on the initially expected payoff, it follows that the frequency of acceptance in $D_2$ is lower than in $D_1$.

We proceed further with the analysis of player A’s behavior. Since A is selfish, his decision utility is only affected by his material payoff, i.e., $u^{D_k}_a(z) = \pi_a(z)$ for every terminal history $z \in Z$ and payoff treatment $k$.

In treatment $D_k$, player A chooses the default allocation $(d)$ if and only if

$$\pi_a(d) = m_a \geq \alpha^{D_k} h := \mathbb{E}_{a,D_k}[\pi_a|g],$$

i.e., if and only if

$$\alpha^{D_k} \leq \hat{\alpha} := \frac{m_a}{h}.$$  

Therefore, A’s best reply to her initial first-order belief about B’s behavior is the default allocation if and only if she attaches sufficiently low probability to player B accepting the greedy offer. Note that the threshold $\hat{\alpha}$ does not depend on the treatment, as we have purposely set A’s payoffs ($m_a$) constant across treatments.

To obtain comparative predictions on the frequency of default allocations in the two treatments, we make a further step in the iterative deductive reasoning and assume not only that players are rational but also that A is certain about B’s (subjective) rationality at the beginning of the game. Moreover, we make an across-

---

32 Based on previous studies on the Ultimatum Game, the 50-50 split was repeatedly found to be the modal offer. Intuitively, it is reasonable that moving to a simple binary choice, most subjects would select a 50-50 split against a highly unbalanced alternative. It can be argued that the even split can be perceived as a focal point in bargaining (Janssen 2006). This argument justifies why B-agents in D1 should believe that A chooses the default allocation at least half of the times and so should B-agents in D2.

33 Notice that, as player A’s payoffs are kept constant across payoff treatments, we can suppress the superscript $D_k$ from the payoff function.
treatment comparative assumption about the distributions of A’s second-order beliefs about B’s expectation of the default allocation. Such beliefs are relevant for our analysis as the higher is the probability assigned by player A to B expecting the default allocation, the higher is A’s subjective probability of a rejection and therefore, due to his selfish preferences, the stronger the incentive to make a default allocation. Thus, we assume that the distributions of A’s second-order beliefs in the two payoff treatments satisfy a stochastic dominance relation, which is in line with the relation between B’s beliefs distributions ($\beta^{D2} \geq \beta^{D1}$). For any version of the conditional probability measure $\phi^{Dk}(\cdot|\theta, \delta)$ (k = 1, 2), we assume that the following condition is satisfied (almost everywhere):

$$P_{\phi^{D2}}(\beta^{D2} \leq \beta|\theta, \delta) \leq P_{\phi^{D1}}(\beta^{D1} \leq \beta|\theta, \delta) \text{ for all } \beta \in [0, 1],$$

(4)
i.e., A-subjects in D2 assign higher probability to B expecting the default allocation than A-subjects in D1.

Therefore, the second step of elimination of non-best replies relies on the following assumption.

**Assumption 2.** (i) Player A is certain that B is subjectively rational and (ii) the distributions of A’s second-order beliefs in D1 and D2 satisfy condition (4).

Consider an individual in role A with personal features ($\alpha, \phi$). Given that he is certain of B’s rationality, according to the strategic reasoning analyzed above (one-shot deviation principle and intrapersonal equilibrium), we obtain his first-order belief on conditional acceptance of the greedy offer as the expected acceptance rate calculated according to his personal second-order beliefs:

$$\alpha = E_{\phi} [\Gamma^{pk}] = \int \Gamma^{pk}(\beta, \theta, \delta) \phi(d\beta, d\theta, d\delta).$$

Under Assumption 2, we can derive a prediction about player A’s first-order beliefs about B’s behavior. Specifically, A-subjects in D1 are expected to attach higher probability to B accepting the greedy offer than A-subjects in D2, as stated in the following lemma.

**Lemma 2.** Under Assumption 2, A’s expectation that B accepts the greedy offer is lower in treatment D2 than in treatment D1, that is,

$$\alpha^{D1} \geq \alpha^{D2}.$$
Proposition 3. The frequency of the default allocation in treatment D2 is higher than in treatment D1, i.e.,

\[ P(a^D2 = d) \geq P(a^D1 = d). \]

As explained above, the predictions that we have derived so far about players’ behavior across payoff treatments apply only to the direct method. Indeed, according to BDS theory, observing the realization of an unexpected event—the greedy offer, in our case—is a necessary condition for experiencing frustration. Thus, in the strategy method we should expect no variation in the frequency of acceptance of the greedy offer due to the payoff increase. In turn, differently from what stated in Lemma 2, there is no reason why player A, who is certain of B’s subjective rationality, should attach a different probability to acceptance of the greedy offer in the two treatments. Thus, also A’s behavior in the strategy method is expected to be independent of the payoff treatment.

However, as explained in Section 3, while the frustration and anger component of player B’s decision utility does not operate in the strategy method, the inequity aversion component is at work, its effect being invariant to the payoff treatment. Both in S1 and S2, player B may dislike the unequal outcome that follows his decision of acceptance. Thus, in both payoff treatments, he will accept the greedy offer if and only if his sensitivity to inequity aversion is low enough \( \delta < \frac{\ell}{h-\ell} \). Therefore, we can expect that the frequency of acceptance of the greedy offer is lower than 100% but it should not vary across S1 and S2. As player A is certain of B’s rationality, both in S1 and S2 he will anticipate that B may reject the greedy offer. Thus, we can expect in turn that the frequency of default allocations is positive in both payoff treatments. Moreover, the frequency of default allocations should be the same in the two treatments since, having similar expectations about player B’s response, A-subjects in S1 will behave similarly to A-subjects in S2.

We can now make comparisons about B’s behavior across methods of play. Keeping fixed the payoff treatment, the method of play is expected to influence the distribution of actions for the following reason. Both in the direct and in the strategy method the inequity aversion component is present and may justify a rejection. However, in the direct method, when frustration and B’s sensitivity to anger are positive, the psychological component can decrease even further B’s willingness to accept the greedy offer. This can be immediately understood by comparing the conditions for rejecting the greedy offer in the two methods of play (given the payoff treatment), as the condition in the strategy method—i.e., \( \delta > \frac{\ell}{h-\ell} \)—is tighter than in the direct method—i.e., \( \delta > \frac{\ell-h\theta \max\{0,\beta m_b-\ell\}}{h-\ell} \).
Proposition 4. For all payoff treatments, the frequency of acceptance of the greedy offer in the strategy method is higher than the conditional frequency of acceptance of the greedy offer in the direct method, i.e., for $k = 1, 2$,

$$
\mathbb{P}(a_b^{Sk} = y) \geq \mathbb{P}(a_b^{Dk} = y | a_a^{Dk} = g). \tag{34}
$$

Under Assumption 2, we can establish a relation between player A’s first-order beliefs about B’s behavior in the two methods of play (for any fixed payoff treatment $k$): A is expected to attach higher probability to player B accepting the greedy offer in the strategy method than in the direct method, i.e., $\alpha^{Sk} \geq \alpha^{Dk}$ ($k = 1, 2$). Thus, player A has clearly more incentives to chose the default allocation in Dk than in Sk, that implies the following result.

Proposition 5. For all payoff treatments, the frequency of the default allocation in the direct method is higher than in the strategy method, i.e., for $k = 1, 2$,

$$
\mathbb{P}(a_a^{Dk} = d) \geq \mathbb{P}(a_a^{Sk} = d).
$$

In a nutshell, the theoretical model and the experimental design allow comparisons of A and B-subjects’ behavior across payoff treatments, given the method of play (Proposition 2 and 3), and across methods of play, given the payoff treatment (Proposition 4 and 5). Summing up, both the payoff increase for B in the default allocation and the direct method should decrease acceptance, but while the payoff increase affects behavior only in the direct method, changing the method of play should affect behavior in both payoff conditions. Consequently, we expect a positive interaction between the negative (positive) effect of the payoff increase and the negative effect of the direct method on the acceptance rate (on the default allocation rate).

5.1 Experimental Hypotheses

The behavioral predictions derived above deliver our experimental hypotheses across treatments. The first set of hypotheses regards the comparison of subjects’ beliefs and behavior across payoff treatments. We will test preliminarily the auxiliary assumption that the distributions of B-subjects’ first-order beliefs satisfy the assumption $\beta^{D2} \geq \beta^{D1}$.

\footnote{Henceforth, to ease the exposition, we will use the wording “frequency of acceptance” for both methods of play meaning the absolute frequency when we refer to the strategy method and the conditional one when we refer to the direct method.}
**Hypothesis 1.1.** *The initially expected payoff of B-subjects in D2 is larger than in D1.*

If this is the case, the larger payoff from the default allocation makes B-subjects’ initially expected payoff larger in D2 than in D1, and we can thus expect a lower probability of accepting the greedy offer in D2 than in D1, due to frustration and anger. Therefore, we can test the main prediction of Proposition 2. Since, as discussed in Section 5, the payoff increase should not have any effect in the strategy method, we will test the following hypothesis.

**Hypothesis 1.2.** *The frequency of acceptance of the greedy offer is lower in D2 than in D1, but it does not differ across S2 and S1.*

For what concerns A-subjects’ beliefs and behavior across treatments, we start by testing whether the implication of Assumption 2 stated in Lemma 2—i.e., that A-subjects in treatment D2 believe that rejections are more likely than A-subjects in treatment D1—finds support in our data. Notice that since A-subjects are assumed to know the utility function of subjects playing in role B—and thus that in the strategy method the psychological component is absent—there should not be any difference in their expectations of acceptance across the two payoff treatments in this condition. We can thus formulate the following experimental hypothesis.

**Hypothesis 1.3.** *A-subjects’ subjective probability of acceptance is lower in D2 than in D1, but it does not differ across S2 and S1.*

Next, we test the prediction of Proposition 3 that relies on A-subjects’ selfishness and their expectations about B-subjects’ behavior.

**Hypothesis 1.4.** *The frequency of the default allocation is higher in D2 than in D1, but it does not differ across S2 and S1.*

The second set of hypotheses relates the comparison of subjects’ behavior across methods of play, as predicted by Proposition 4 and 5.

**Hypothesis 2.1.** *For all payoff treatments, the frequency of acceptance of the greedy offer is lower in the direct than in the strategy method.*

**Hypothesis 2.2.** *For all payoff treatments, the frequency of the default allocation is higher in the direct than in the strategy method.*

---

35Since we do not elicit second-order beliefs, we cannot directly test whether the assumption that A-subjects in treatment D2 expect higher values of B-subjects first-order beliefs about the default allocation than A-subjects in treatment D1. Yet, we can test its implication on A-subjects’ first-order beliefs about acceptance.
Besides the comparison of subjects’ beliefs and behavior across treatments, we can test some further predictions (given the method of play and the payoff treatment) implied by our theoretical model. The analysis of player B’s behavior of Section 3—specifically, the assumptions about B’s subjective rationality and his preferences—delivers the following hypothesis:

**Hypothesis 3.1.** *In the direct but not in the strategy method, the higher the initial expectation of a default allocation, the lower the probability of accepting a greedy offer.*

The second hypothesis concerns the relation between A-subjects’ beliefs and behavior. We test whether there is a correlation between A-subjects’ expectation of acceptance and the probability of the default allocation. While the correlation of Hypothesis 3.1 is expected to hold in the direct method only, due to the psychological preferences of B-subjects, the second is expected to hold in both methods of play.

**Hypothesis 3.2.** *In both methods of play, the higher the expectation of acceptance, the higher the probability of making a greedy offer.*

Notice that evidence in favor of Hypothesis 3.2 would be compatible with the selfishness assumption that we made, but it does not fully exclude that A-subjects may be affected by some sort of social preferences (e.g., inequity aversion). This last hypothesis, like in general the hypotheses about A-subjects, is less relevant for our purposes as it is not directly implied by the theory of frustration and anger applied to the UMG. Nonetheless, we can use the values of the Aquino morality scale as a proxy for inequity aversion.

### 6 Results

This section presents the experimental results. We start by providing a description of B-subjects’ behavior. Next, we report results for A-subjects’ behavior.

There were 176 B-subjects in total, with 66 facing the strategy method and 110 the direct method. Only 72 out of the 110 B-subjects in the direct method received the greedy offer and did actually make a choice. Thus, we have in total 138 observations of B-subjects’ behavior.

Before examining B-subjects’ behavior, we check whether our auxiliary assumption regarding their initial first-order belief about the default allocation is verified \( (\beta^{D2} \geq \beta^{D1}) \). Table 2 shows the average subjective probabilities of the default allocation.
Allocation of B-subjects prior to choices in the experimental UMG, by payoff treatment and method of play. The table also reports the differences in means with the relative p-value of the tests. Results of both a two-sample t-test of the difference in means and of a Mann-Whitney U test (henceforth, MW) indicate that there is no significant difference in initial beliefs between D1 and D2, the two estimates being virtually identical. Thus, our auxiliary assumption is verified and B-subjects’ initially expected payoff is larger in D2 than in D1, as stated in Hypothesis 1.1.

Instead, a p-value of 0.06 (MW)—and of 0.04 for the t-test—indicates that there is a significant difference between beliefs in S1 and S2, with higher values in S1. Notice however that our predictions about B-subjects’ behavior do not rely on auxiliary assumptions regarding their initial first-order beliefs in the strategy method, because according to the BDS model psychological preferences do not matter in this method of play.

Table 2: B-subjects’ expectations of the default allocation

<table>
<thead>
<tr>
<th>Mean β</th>
<th>( m^1 )</th>
<th>( m^2 )</th>
<th>Diff.</th>
<th>p-value (t-test)</th>
<th>p-value (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>0.54</td>
<td>0.54</td>
<td>0.00</td>
<td>1.00</td>
<td>0.93</td>
</tr>
<tr>
<td>Strategy</td>
<td>0.62</td>
<td>0.48</td>
<td>0.15</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>Diff.</td>
<td>-0.09</td>
<td>0.06</td>
<td></td>
<td>0.14</td>
<td>0.10</td>
</tr>
<tr>
<td>p-value (t-test)</td>
<td>0.14</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value (MW)</td>
<td>0.10</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: means of B-subjects’ subjective probability of the default allocation by payoff treatment and method of play; differences across treatments and p-values of a two-sided t-test of the differences in means and of a Mann-Whitney U test are displayed.

Figure 4 illustrates the share of B-subjects accepting the greedy offer by treatment. Overall, the average frequency of acceptance of the greedy offer is 65%. The observed frequency of acceptance is lower in the direct than in the strategy method (MW p-value 0.008; 138 observations) and it is similar across payoff treatments both in the direct (MW p-value 0.65; 72 observations) and in the strategy method (MW p-value 0.57; 66 observations). This evidence seems to provide support to Hypothesis 2.1 (lower frequency of acceptance in the direct method), but it is at odds with Hypothesis 1.2 (higher frequency of acceptance in D1 than in D2, but constant across S1 and S2).

Next, we analyze the effect of B-subjects’ initial expectations on the probability of accepting a greedy offer. Table 3 reports the means of B-subjects’ subjective probabilities of a default allocation conditional on accepting or rejecting a greedy offer (by method of play). In line with our predictions, there is a significant negative
correlation between B-subjects’ initial expectations and the probability of accepting a greedy offer in the direct method. B-subjects rejecting the greedy offer expected on average that the default allocation was 20% more likely than those who accepted it. We also find mild evidence of a negative correlation between B-subjects’ initial expectations and the probability of accepting a greedy offer in the strategy method. The last correlation was unexpected though, thus Hypothesis 3.1 (positive correlation between initial expectations and rejections in the direct method and no correlation at all in the strategy method) is only partially verified.

Table 3: Initial expectations and rejections

<table>
<thead>
<tr>
<th>Mean $\beta$</th>
<th>Accept</th>
<th>Reject</th>
<th>Diff.</th>
<th>p-value (t-test)</th>
<th>p-value (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>0.42</td>
<td>0.62</td>
<td>-0.20</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Strategy</td>
<td>0.51</td>
<td>0.67</td>
<td>-0.16</td>
<td>0.05</td>
<td>0.08</td>
</tr>
<tr>
<td>Diff.</td>
<td>-0.09</td>
<td>-0.06</td>
<td></td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td>p-value (t-test)</td>
<td>0.16</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value (MW)</td>
<td>0.18</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: means of B-subjects’ subjective probability of the default allocation conditional on accepting and rejecting the greedy offer, by method of play; differences across conditions and p-values of a two-sided t-test of the differences in means and of a Mann-Whitney U test are displayed.

In order to estimate the effects of our treatments on the probability of accepting
a greedy offer, we use the following probit model:

\[
P(y_i = 1 | \text{Direct}_i, \text{Payoff Increase}_i, x_i) = \Phi(\lambda_0 + \lambda_1 \text{Direct}_i + \lambda_2 \text{Payoff Increase}_i + \lambda_3 \text{Direct}_i \times \text{Payoff Increase}_i + \lambda_4' x_i),
\]

where \( \Phi \) is the standard normal cumulative distribution function, \( y_i \), with \( i = 1, \ldots, 138 \), is a dummy that takes value 1 when subject \( i \) accepts the greedy offer, \( \text{Direct}_i \) is a dummy that takes value 1 (0) in the direct method (strategy method), \( \text{Payoff Increase}_i \) is a dummy that takes value 1 (0) in payoff treatment \( m_2 \) (\( m_1 \)) and \( x_i \) is a vector of controls that may or may not include (depending on the specification) \( i \)'s beliefs (on a scale from 0 to 100) and their interaction with \( \text{Direct}_i \), Age, Gender (dummy that takes value 1 if female) and Trait Anger (as elicited from STAXI-II)\(^{37}\).

Table 4 reports the average marginal effects of the treatment variables on the probability of acceptance as implied by different specifications of the probit estimation. The upper part of column I reports the average marginal effects of the two treatments and their interaction (\( \text{Direct} \times \text{Payoff Increase} \)), which is computed as the difference between the marginal effects of Payoff Increase in the two methods of play (reported in the lower part of the table). Column II reports results from an alternative specification of equation (5) that includes Belief—i.e., \( i \)'s expectation of the default allocation—and its interaction with Direct as explanatory variables. Specifically, Belief \( \times \) Direct refers to the difference between the marginal effects of beliefs in the two methods of play. Column III reports the same estimates for a specification of equation (5) that includes Trait Anger, Age and Gender as controls.

In line with Hypothesis 2.1, the estimated probability of accepting a greedy offer is lower in the direct than in the strategy method, as indicated by a negative and significant average marginal effect of Direct. Against Hypothesis 1.2, the estimates do not indicate a negative effect of the payoff increase on the probability of accepting the greedy offer, the marginal effect of Payoff Increase being not significant in either method of play.

For what regards Hypothesis 3.1, B-subjects’ first-order beliefs have a significant negative effect on the probability of acceptance in both methods of play. Yet, the negative impact of beliefs is amplified in the direct method: against an increase

\(^{37}\)Trait Anger measures the individual’s general disposition toward anger as a stable trait, differently from State Anger that measures the individual’s current emotional state. We do not use State Anger as a control since it is measured at the end of the experiment, after all choices are made and payoffs are realized, and thus may not correctly identify B-subjects’ emotions at the information set following the greedy offer. Note that the Trait Anger, as expected, does not correlate with the final payoff (p-value 0.793).
Table 4: B-subjects’ behavior

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>-0.217***</td>
<td>-0.246***</td>
<td>-0.249***</td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.072)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>Payoff Increase</td>
<td>0.057</td>
<td>0.058</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.074)</td>
<td>(0.074)</td>
</tr>
<tr>
<td>Direct × Payoff Increase</td>
<td>-0.007</td>
<td>0.070</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>(0.158)</td>
<td>(0.145)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Belief</td>
<td>-0.006***</td>
<td>-0.006***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Belief × Direct</td>
<td>-0.004</td>
<td>-0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>Trait Anger</td>
<td>-0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.042*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.064</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payoff Increase in direct method</td>
<td>0.053</td>
<td>0.091</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>(0.117)</td>
<td>(0.104)</td>
<td>(0.104)</td>
</tr>
<tr>
<td>Payoff Increase in strategy method</td>
<td>0.061</td>
<td>0.021</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.101)</td>
<td>(0.099)</td>
</tr>
<tr>
<td>Observations</td>
<td>138</td>
<td>138</td>
<td>138</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is a dummy that indicates whether i accepts the greedy offer; belief the table reports average marginal effects estimated from probit models; for the treatments interaction, we report the differential marginal effects of Payoff Increase when Direct changes from 0 to 1 (conditional marginal effects are reported in the lower part of the table); for the interaction of Belief with Direct, we report the differential marginal effect of Belief when Direct changes from 0 to 1; standard errors in parentheses; significance levels are: * p < 0.1; ** p < 0.05; *** p < 0.01.
of 10% in the expectation of the default allocation, the probability of acceptance decreases (on average) of approximately 6% points in the strategy method and of 9% in the direct method. The negative difference of 4% points between the marginal effect of belief in the direct and in the strategy method displays a p-value of 0.12. The negative interaction between beliefs and direct method is clearly illustrated in Figure 5 that shows the predicted probability of acceptance against beliefs from a probit model with beliefs, treatments dummies and their interaction as regressors. Beliefs have a negative effect on the probability of acceptance not only in the direct method but, contrary to our predictions, also in the strategy method. The latter effect is smaller though.

Figure 5: Predicted probability of acceptance against B’s first-order beliefs

Among controls, Trait Anger has a (mildly significant) negative effect (p-value: 0.13) on the probability of accepting a greedy offer. In line with the theory of frustration and anger, the higher the sensitivity to anger, the lower it is the probability of acceptance.

Let us now describe our experimental findings about A-subjects’ behavior and beliefs. We have in total 176 observations for A-subjects’ (66 for Strategy and 110 for Direct). The overall frequency of default allocations is 36%. Figure 6 illustrates the share of A-subjects making the default allocation by treatment. The figure suggests that neither the method of play (MW: p-value 0.65; 176 observations) nor the payoff treatment in the direct method (MW p-value 0.69; 110 observations) and in the strategy method (MW p-value 0.45; 66 observations) have a significant effect on the frequency of the default allocation.

In line with Hypothesis 3.2, we find a negative correlation between the probability of choosing the default allocation and the expectation of acceptance (Spearman
correlation coefficient equal to -0.34, p-value 0.00). Table 5 displays the t-tests of difference in means of the beliefs distribution for A-subjects across payoff treatments and methods of play. Both the t-test and the MW test indicate no significant difference in A-subjects’ beliefs across payoff treatments (for fixed method of play). That is, we do not find evidence in favor of Hypothesis 1.3—i.e., that A-subjects’ expectation of acceptance is higher in D1 than in D2. Also, in line with B-subjects’ behavior, A-subjects’ expectations of acceptance are higher in the strategy than in the direct method, significantly so in payoff treatment $m_1$. Yet, as shown in Table 3, notwithstanding these first-order beliefs—and the negative correlation between expectation of acceptance and probability of the default allocation—A-subjects are not more likely to choose the default allocation in the direct method.

We estimate the effects of the treatments on A-subjects’ behavior by considering a probit estimate of the probability of choosing the default allocation with Direct, Payoff Increase and their interaction as explanatory variables. As for B-subjects, we consider alternative specifications of the baseline probit model. In specification II we also consider the marginal effect of A-subjects’ first-order beliefs about their co-player’s move. In specification III we control for individual observable charac-

---

38Given our payoff structure, a selfish player A should make the greedy offer if his expectation of acceptance is higher than 0.57. The observed lowest average expectation is higher than this threshold (0.70 in D1). Yet, if we limit the attention to those A-subjects having an expectation of acceptance higher than 0.57, not all of them (73%) play the greedy offer. Thus, there is a share of A-subjects whose behavior may be affected by a concern for the material payoff of the opponent.

39Interestingly, A-subjects have a correct estimate of the probability of acceptance in S1 (0.82 versus the actual share which is 0.73) and S2 (0.79 versus 0.79), while they overestimate the probability of acceptance in D1 (0.70 versus 0.51) and D2 (0.73 versus 0.57).

40Unlike Table 4, we do not include the interaction Belief $\times$ Direct in specification II and III.
Table 5: A-subjects’ expectation of acceptance

<table>
<thead>
<tr>
<th></th>
<th>Mean $\alpha$</th>
<th>$m_b^1$</th>
<th>$m_b^2$</th>
<th>Diff.</th>
<th>p-value (t-test)</th>
<th>p-value (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>0.70</td>
<td>0.73</td>
<td>-0.03</td>
<td>0.59</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>Strategy</td>
<td>0.82</td>
<td>0.79</td>
<td>0.03</td>
<td>0.58</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>-0.12</td>
<td>-0.06</td>
<td></td>
<td></td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>p-value (t-test)</td>
<td>0.03</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value (MW)</td>
<td>0.01</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Means of A-subjects’ subjective probability of acceptance, by payoff treatment and method of play; differences across treatments and p-values of a two-sided t-test of the differences in means and of a Mann-Whitney U test are displayed.

characteristics (Age and Gender) and the subject’s degree of morality—as elicited by the Aquino questionnaire—and the subject’s degree of guilt aversion—as elicited by the GASP questionnaire.

The upper part of Table 6 reports the average marginal effects of the treatments dummies and the effect of their interaction, which is computed as the difference between the marginal effects of Payoff Increase in the two methods of play (reported in the lower part of the table) on the probability of choosing the default allocation, as implied by the probit estimate. Estimated marginal effects indicate that, contrary to Hypothesis 1.4, the payoff increase does not have any significant effect on the probability of the default allocation in either method of play. Also, differently from B-subjects and at odds with Hypothesis 2.2, A-subjects’ behavior seems not to be affected by the method of play. Instead, in line with Hypothesis 3.2, probit estimates confirm that A-subjects’ first-order beliefs have a significant negative effect on the probability of acceptance. Morality and Guilt aversion seem not to affect A-subjects’ behavior. The estimated marginal effect of Gender indicates that females are more likely to choose the default allocation than males.

Summarizing, our experimental results so far are as follows.

• Result 1.1: The average initially expected payoff of B-subjects in D2 is larger than in D1 (Hypothesis 1.1 is verified).

• Result 1.2: The frequency of acceptance of the greedy offer is not significantly different both across S2 and S1 and across D1 and D2 (Hypothesis 1.2 is not verified).

• Result 1.3: A-subjects’ average subjective probability of acceptance is not
Table 6: A-subjects’ behavior

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>-0.033</td>
<td>-0.100</td>
<td>-0.068</td>
</tr>
<tr>
<td></td>
<td>(0.075)</td>
<td>(0.068)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>Payoff Increase</td>
<td>-0.057</td>
<td>-0.052</td>
<td>-0.033</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.066)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Direct × Payoff Increase</td>
<td>0.055</td>
<td>0.101</td>
<td>0.101</td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td>(0.136)</td>
<td>(0.131)</td>
</tr>
<tr>
<td>Belief</td>
<td>-0.008***</td>
<td>-0.007***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Morality</td>
<td></td>
<td>0.051</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.060)</td>
<td></td>
</tr>
<tr>
<td>Guilt aversion</td>
<td></td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.031)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>0.189***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.063)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.022)</td>
</tr>
<tr>
<td>Payoff Increase in direct method</td>
<td>-0.036</td>
<td>-0.014</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.091)</td>
<td>(0.080)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>Payoff Increase in strategy method</td>
<td>-0.091</td>
<td>-0.115</td>
<td>-0.095</td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td>(0.110)</td>
<td>(0.105)</td>
</tr>
<tr>
<td>Observations</td>
<td>176</td>
<td>176</td>
<td>176</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is a dummy that indicates whether i chooses the default allocation; the table reports average marginal effects estimated from probit models; for the treatments interaction, we report the differential marginal effects of Payoff Increase when Direct changes from 0 to 1 (conditional marginal effects are reported in the lower part of the table); standard errors in parentheses; significance levels are: * p < 0.1; ** p < 0.05; *** p < 0.01.
significantly different across payoff treatments, in both methods of play (Hypothesis 1.3 is not verified).

• **Result 1.4:** The frequency of the default allocation is not significantly different across payoff treatments, in both methods of play (Hypothesis 1.4 is not verified).

• **Result 2.1:** The frequency of acceptance of the greedy offer is significantly lower in the direct method than in the strategy method (Hypothesis 2.1 is verified).

• **Result 2.2:** The frequency of the default allocation is not significantly different across methods of play (Hypothesis 2.2 is not verified).

• **Result 3.1:** There is a significant negative correlation between the initial expectations of a default allocation and the probability of accepting the greedy offer in the direct method, as predicted by Hypothesis 3.1, but we also find mild evidence of such correlation in the strategy method. Yet, the negative effect of initial expectations on the probability of acceptance is larger in the direct method than in the strategy method.

• **Result 3.2:** In both methods of play, there is a significant negative correlation between the expectation of acceptance and the probability of choosing the default allocation (Hypothesis 3.2 is verified).

In a nutshell, among the hypotheses about the comparison of behavior across treatments, only those about B-subjects’ behavior across methods of play (Hypothesis 2.1) is verified. Instead, the two experimental hypotheses about A and B-subjects’ beliefs and behavior within treatments (Hypotheses 3.1 and 3.2) are both confirmed by the data. Certainly, the payoff manipulation does not seem to have the expected effect either on the frequency of acceptance or of default allocations. We discuss possible explanations of this result in Section 7.

### 7 Discussion

Our experimental results show that the manipulation of player B’s payoff from the default allocation does not have a significant effect on B-subjects’ decision to accept or reject the greedy offer. However, in line with the theory of frustration and anger, we find a significant positive effect of initial beliefs on rejections. Given that B-subjects’ beliefs are on average similar in payoff treatments D2 and D1, and the
payoff manipulation is such that B’s payoff in the default allocation is larger in D2 than in D1, B-subjects’ initially expected payoff should be on average larger in D2 than in D1. Yet, according to the data, rejections are not more frequent in D2 than in D1. The evidence suggests that there is another force that contrasts and neutralizes the predicted effect of the increase of B-subjects’ initial expectations on rejections.

A possible explanation relies on the fact that, while in D1 the default allocation is an equal split, in D2 it gives more to B. The default allocation is in both cases the most upright choice. Indeed, choosing the greedy offer cannot be justified by any sort of entitlement of A-subjects to obtain a larger payoff than B-subjects. Yet, due to the equal-split default, choosing the greedy offer in D1 may be perceived by B-subjects as a more severe violation of a moral standard than in D2. Indeed, it is plausible that the normative strength of the default allocation decreases as inequity increases, so that the deviation from a moral standard through the greedy offer is maximal when the outside option is most equitable, i.e., in D1.

Clearly, when the perception of the moral outrage is stronger, there is less room for the frustration and anger motivation triggered by goal obstruction. In other words, some B-subjects might blame more A-subjects for being unfair than for diminishing their expected material payoff, and may thus blame A-subjects more in D1 than in D2—i.e., when their own payoff from the default allocation is smaller but moral outrage is stronger.

The BDS definition of frustration and anger does not account for B-subjects’ reaction to the violation of a moral standard. Actually, B-subjects’ sensitivity to moral outrage is a potential confounding variable as it may correlate with their initial expectations: subjects who attach higher value to the equal split might also expect the default allocation with higher probability. However, if this mechanism were pervasive, we would observe that, contrary to our findings, in D1 B-subjects’ expectations of the default allocation are higher than in D2.

We might imagine an extension of the model that incorporates two definitions of anger: anger provoked by goal blockage that triggers a rejection to retaliate against the harm-doer, and anger as moral outrage that triggers a rejection to restore the violated moral standard (Frijda 1986). In our context, both types of negative

---

41Notice that A-subjects’ violation of a moral standard per se may induce B-subjects to reject the greedy offer even when they do not experience frustration due to unfulfilled expectations.

42Yet, this correlation is not obvious: if an individual is sensitive to the violation of some moral standard, he does not necessarily expects that others will not violate it. In fact, the concept of moral standard we relate to is weaker than that of social norm: a social norm is commonly believed to be spread in a society and it is expected that most individuals will comply with it.

43Notice moreover that against the hypothesis of the equal split as a social norm, we also find that A-subjects’ expectations of acceptance are not lower in D1 than in D2.
emotional responses can be triggered by a greedy offer. There might be individuals influenced by both types of anger, individuals affected by only one of the two, and individuals affected by none of them. Certainly, while the first type of anger is more likely to be triggered in D2 than in D1, the opposite holds for the second type of anger.

In Appendix B we report findings about A and B-subjects’ behavior by gender. Since our experiment was not set out to test for gender differences, any comparison across the females’ and males’ sample need to be considered with due caution. The number of females and males subjects differ—with a large majority of males—and we did not control for the gender distributions within sessions. Yet, we consider worthwhile to discuss these results as—even if only suggestive of the possible phenomena at work—differences of behavior across genders may help interpret the absence of a payoff treatment effect. In a nutshell, while male B-subjects display a higher frequency of rejections in D2 than in D1 (not significant though), female B-subjects display a higher frequency of rejections in D1 than in D2. This evidence suggests that female B-subjects may be especially affected by anger as moral outrage. Instead, males seem more sensitive to anger due to goal blockage. Moreover, evidence about A-subjects’ behavior by gender suggests that female (male) subjects internalize the response of their female (male) counterpart to the greedy offer. Thus, the opposite effect that the payoff increase has on B-subjects’ behavior across genders (in the direct method) is not likely to be a coincidence and it may be due to the fact that females and males appraise the same event of receiving the greedy offer in a different way.

According to Social Role Theory (Eagly and Wood 2011), gender differences in behavior derive from the different social roles that females and males have in a society. Given that one gets angry when his/her major concerns are at stake (Frijda 1986) and the engagement in a social role obviously determines one’s concerns, Social Role Theory might contribute to explain why anger is differently generated for men and women. On the one hand, males were originally the primary family providers, more concerned with power-based emotions and personal evaluation achievement. Thus, they may still be more individually-oriented and perceive the greedy offer as a blockage for their personal achievement. On the other hand, females were originally the primary caretakers, more concerned with interpersonal relations and relational consequences. Thus, they may have evolved a more socially oriented attitude.

Two experiments have studied gender differences in the Ultimatum Game (Eckel

---

44In particular, according to Social Role Theory, physical sex differences have determined a sexual division of labor, which is observed by the society and internalized through gender role beliefs, i.e., gender specific expectations and normative standards.
and Grossman 2001, Solnick 2001) finding mixed evidence about gender differences in responders’ behavior. Yet, evidence from several experimental studies supports the view that women care more about equality (Eckel and Grossman 1998, Andreoni and Vesterlund 2001, Güth et al. 2007). Such evidence is consistent with the hypothesis that females have a high moral standard and—when this is not attained—they express a firm negative emotional response (moral outrage).

Responders’ behavior could be also explained by negative reciprocity, an alternative motivation that can trigger aggressive behavior (Rabin 1993, Dufwenberg and Kirchsteiger 2004, Falk and Fischbacher 2006): a rejection may be the reaction to the perceived unkindness of the proposer associated to the greedy offer. Reciprocity à la Dufwenberg and Kirchsteiger (2004) delivers behavioral predictions that are in line with some of our hypothesized treatment effects. In particular, reciprocity predicts a higher rejection rate in the direct method. Indeed, the greedy offer is always perceived as unkind in the direct method, while in the strategy method the proposer’s unkindness depends on his first-order beliefs, and thus the responder’s willingness to reciprocate may be attenuated. Moreover, increasing the responder’s payoff in the default allocation mechanically increases the proposer’s unkindness, leading to more rejections. Yet, unlike anger, this effects should kick in also in the strategy method.

Importantly, reciprocity and anger deliver different predictions regarding the correlation between the responder’s behavior and beliefs. In the direct method, the willingness to reject under reciprocity does not depend on the responder’s first-order beliefs, but only on the perceived intentions of the proposer, captured by the responder’s second-order belief: the more B believes that A expects him to accept, the more kind he believes the proposer is, the less B is willing to reject the greedy offer. In the strategy method, the responder’s decision depends not only on perceived intentions but also on his first-order beliefs: the more B expects the default allocation, the more kind he believes the proposer is, the less B is willing to reject the greedy offer. Thus, the positive correlation between responders’ first-order beliefs

---

Notes:

45 Croson and Gneezy (2009), who survey the experimental literature on gender differences in social dilemmas, argue the different findings of the two studies are probably due to the experimental context: while in Eckel and Grossman (2001) subjects interact face to face and with the direct method, in Solnick subjects interact anonymously and with the strategy method.

46 Interestingly, female subjects of our experiment scored higher than males in the Aquino questionnaire on moral identity.

47 Unlike anger, which cannot affect behavior at the root of the game, reciprocity can shape decisions at the empty history as well. To better compare results with our analysis, here we consider role-dependent preferences, with the first-mover having standard preferences.

48 In details, if a B-player expects the greedy offer with more than 50% probability—as found on average in the data—he always accepts; otherwise, the more he expects the default allocation, the less he accepts.
and rejections in the direct method can be explained by frustration and anger but not by reciprocity. The mildly significant positive correlation between responders’ first-order beliefs and rejections in the strategy method is at odds with our model that predicts zero correlation, but it is even more in contrast with reciprocity that actually predicts a negative correlation. We discussed in previous sections how cross-disciplinary findings support the relevance of initial expectations in negotiations and this is the main reason why we favor anger due to unfulfilled expectations as a motivation of rejections.\footnote{Section 6 of van Damme et al. (2014) by Dufwenberg and Kirchsteiger discusses the role of reciprocity in the UG using equilibrium analysis to explain the evidence of Blount (1995). These results can be explained by the difference between two functional forms of BDS: simple anger and anger from blaming behavior. See also Section 4 of Battigalli et al. (2019b) for a comparison between reciprocity and anger.} However, one can argue that different social emotions shape the strategic interaction during a negotiation and negative reciprocity and anger may simultaneously affect players’ decisions. We do not dismiss this hypothesis, which deserves further investigation.

8 Conclusion

In this paper we tested experimentally the theory of frustration and anger in the strategic context of the Ultimatum Minigame. The evidence that we find is mixed. From a theoretical perspective, a first important contribution of this paper consists of deriving testable behavioral predictions without assuming either complete information, or equilibrium play. It is only assumed that incompletely informed players are rational, confident in others’ rationality and their beliefs satisfy some plausible assumptions.

The experimental analysis delivers an important result in support of the theory of frustration and anger: responders’ behavior depends on initial expectations. Specifically, we find that the higher the responder’s expectation of the default allocation, the more likely it is that he rejects the greedy offer. This is in line with the theory of frustration and anger that assumes that the responder’s decision utility depends on his initially expected payoff: the larger the responder’s expected payoff at the root of the game, the higher his frustration when observing the greedy offer and thus, the more willing he is to reject it.

We have measured the impact of initial expectations on behavior by eliciting personal beliefs at the root of the game. This innovates on previous studies which considered the impact on behavior of expectations that are induced by the experimenter, by providing information on average behavior of proposers in previous sessions or in other matches in the same session. Such procedure may induce a social norm (play-
ing the default allocation) and make it difficult to disentangle whether rejections are an emotional reaction to the observed deviation from this norm (the greedy offer) or to diminished personal expectations. Instead, thanks to beliefs elicitation, we can safely conclude that initial personal beliefs influence responders’ behavior.

Another important result, related to the first, regards the method of play. Rejections are more frequent in the direct than in the strategy method and in the direct method the negative effect of initial expectations on the probability of accepting the greedy offer is stronger than in the strategy method. The differential effect of beliefs in the two methods of play provides additional support to the theory of frustration and anger. Indeed, not only initial beliefs matter, but the fact that they matter more when responders can actually observe the greedy offer before making their choice confirms that frustration is a credible explanation for rejections. These results suggest that using the direct method is more appropriate when observing previous moves is likely to induce emotional reactions.

Finally, we have extensively discussed the possible reasons why the payoff manipulation has no effect on aggregate behavior. Such discussion provides important insights for future research. In particular, our analysis reveals that there might be alternative models of frustration and aggressive behavior, in which gender may play a significant role.
Appendix A

Derivation of intrapersonal equilibrium correspondence $\Gamma$. B expects to accept the greedy offer, i.e., $\gamma = 1$, if

$$\ell - h\theta F^{D,m_b}(\beta, 1) - \delta(h - \ell) > 0,$$

that is,

$$\delta < \delta^{D,m_b}(\beta, 1, \theta) = \frac{\ell - h\theta (\beta m_b + (1 - \beta) \ell - \ell)}{h - \ell} = \frac{\ell - h\theta \beta(m_b - \ell)}{h - \ell}.$$

B instead expects to reject the greedy offer, i.e., $\gamma = 0$, if

$$\ell - h\theta F^{D,m_b}(\beta, 0) - \delta(h - \ell) < 0,$$

that is,

$$\delta > \delta^{D,m_b}(\beta, 0, \theta) = \frac{\ell - h\theta \max\{0, \beta m_b - \ell\}}{h - \ell}.$$

Since $\max\{0, \beta m_b - \ell\} \leq \beta (m_b - \ell)$, we obtain the inequality between thresholds:

$$\delta^{D,m_b}(\beta, 1, \theta) = \frac{\ell - h\theta \beta(m_b - \ell)}{h - \ell} \leq \frac{\ell - h\theta \max\{0, \beta m_b - \ell\}}{h - \ell} = \delta^{D,m_b}(\beta, 0, \theta).$$

Finally, if B expects to accept with probability $\gamma \in (0, 1)$, then he has to be indifferent:

$$\ell - h\theta F^{D,m_b}(\beta, \gamma) - \delta(h - \ell) = 0,$$

that is,

$$\delta = \frac{\ell - h\theta F^{D,m_b}(\beta, \gamma)}{h - \ell} \in \left[\delta^{D,m_b}(\beta, 1, \theta), \delta^{D,m_b}(\beta, 0, \theta)\right].$$

In this case we pin down the “rational-expectations” value of $\gamma$ as the solution of the following indifference condition:

$$\ell - h\theta (\beta m_b + (1 - \beta) \gamma \ell - \ell) - \delta(h - \ell) = 0,$$

that is

$$\gamma = \frac{\ell - h\theta (\beta m_b - \ell) - \delta(h - \ell)}{h\theta (1 - \beta) \ell} \text{ if } \theta > 0 \text{ and } 0 < \beta < 1.$$

The acceptance probability is not pinned down by the intrapersonal equilibrium condition if and only if $\delta = \delta^{D,m_b}(\beta, 1, \theta) = \delta^{D,m_b}(\beta, 0, \theta)$. The second equality holds if and only if either $\theta = 0$ (frustration does not matter), or $\beta = 0$ (no frustration), or $\beta = 1$ (the initially expected payoff is independent of $\gamma$).
Proof of Lemma \[\text{Consider } \beta, \beta' \text{ such that } \beta' > \beta. \text{ First, it is easy to check that } \hat{\delta}_{D,m_b}^{D,m_b}(\beta', 1, \theta) \leq \hat{\delta}_{D,m_b}^{D,m_b}(\beta, 1, \theta) \text{ and } \hat{\delta}_{D,m_b}^{D,m_b}(\beta', 0, \theta) \leq \hat{\delta}_{D,m_b}^{D,m_b}(\beta, 0, \theta). \text{ We have to consider the non-trivial case where } \hat{\delta}_{D,m_b}^{D,m_b}(\cdot, 1, \theta) < \hat{\delta}_{D,m_b}^{D,m_b}(\cdot, 0, \theta) \text{ for both } \beta \text{ and } \beta', \text{ and } \delta \text{ falls in the intermediate region: } \delta \in \left[\hat{\delta}_{D,m_b}^{D,m_b}(\beta, 1, \theta), \hat{\delta}_{D,m_b}^{D,m_b}(\beta', 0, \theta)\right]. \text{ We need to show that}\]

\[
\frac{\ell - h\theta (\beta' m_b - \ell) - \delta(h - \ell)}{h\theta(1 - \beta') \ell} \leq \frac{\ell - h\theta (\beta m_b - \ell) - \delta(h - \ell)}{h\theta(1 - \beta) \ell}
\]

\[
\Leftrightarrow (1 - \beta) [\ell - h\theta (\beta' m_b - \ell) - \delta(h - \ell)] \leq (1 - \beta') [\ell - h\theta (\beta m_b - \ell) - \delta(h - \ell)]
\]

\[
\Leftrightarrow (\beta' - \beta) [\ell - h\theta (m_b - \ell) - \delta(h - \ell)] \leq 0.
\]

By assumption \((\beta' - \beta) > 0\). The term in brackets is negative because:

\[
\ell - h\theta (m_b - \ell) - \delta(h - \ell) \leq \ell - h\theta (\beta m_b + (1 - \beta) \ell - \ell) - \delta(h - \ell) \leq 0
\]

where the first inequality follows from \(m_b > \ell\) and the second follows from \(\hat{\delta}_{D,m_b}^{D,m_b}(\beta, 1, \theta) \leq \delta \leq \hat{\delta}_{D,m_b}^{D,m_b}(\beta', 0, \theta);\) indeed,

\[
\delta \geq \hat{\delta}_{D,m_b}^{D,m_b}(\beta, 1, \theta) \Leftrightarrow \delta \geq \frac{\ell - h\theta (\beta m_b + (1 - \beta) \ell - \ell)}{h - \ell}
\]

\[
\delta(h - \ell) \geq \ell - h\theta (\beta m_b + (1 - \beta) \ell - \ell)
\]

\[
\ell - h\theta (\beta m_b + (1 - \beta) \ell - \ell) - \delta(h - \ell) \leq 0,
\]

where latter inequality is implied by the indifference condition.

Proof of Proposition 2. It is easily verified that \(\Gamma\) is decreasing in \(m_b\):

\[
\frac{\partial}{\partial m_b} \Gamma^{D_k}(\beta, \theta, \delta) = \begin{cases} 
-\frac{h\theta \beta}{h\theta(1 - \beta) \ell} & \text{if } \delta \in \left[\hat{\delta}_{D,m_b}^{D,m_b}(\beta, 1, \theta), \hat{\delta}_{D,m_b}^{D,m_b}(\beta, 0, \theta)\right], \hat{\delta}_{D,m_b}^{D,m_b}(\beta, 1, \theta) \neq \hat{\delta}_{D,m_b}^{D,m_b}(\beta, 0, \theta) \\
0 & \text{otherwise.}
\end{cases}
\]

With this, since \(m_b^1 < m_b^2\), for all \((\beta, \theta, \delta)\) it holds that

\[
\Gamma^{D_1}(\beta, \theta, \delta) \geq \Gamma^{D_2}(\beta, \theta, \delta).
\]

Hence

\[
\mathbb{E} [\gamma^{D_1}] = \int \Gamma^{D_1} d\mu^{D_1} \geq \int \Gamma^{D_2} d\mu^{D_1}.
\]

Consider any version of the conditional probability measures \(\mu^{D_k}(\cdot|\theta, \delta)\) \((k=1,2)\).
Since, due to Lemma 1, $\Gamma_D^2$ is decreasing in $\beta$, and we assumed $\beta^D_2 \geq \beta^D_1$, it follows that

$$\int \Gamma_D^2 (\beta, \theta, \delta) \mu^D_1 (d\beta|\theta, \delta) \geq \int \Gamma_D^2 (\beta, \theta, \delta) \mu^D_2 (d\beta|\theta, \delta).$$

Therefore, taking into account that the marginal distribution of $(\theta, \delta)$ is the same in the two treatments because they are personal traits (hence exogenous), we obtain

$$\int \Gamma_D^2 d\mu^D_1 = \int \left( \int \Gamma_D^2 (\beta, \theta, \delta) \mu^D_1 (d\beta|\theta, \delta) \right) \text{marg}_{\theta, \delta} \mu^D_1 (d\theta, d\delta) \geq \int \left( \int \Gamma_D^2 (\beta, \theta, \delta) \mu^D_2 (d\beta|\theta, \delta) \right) \text{marg}_{\theta, \delta} \mu^D_2 (d\theta, d\delta) = \int \Gamma_D^2 d\mu^D_2.$$

Hence

$$E \left[ \gamma_D^1 \right] \geq \int \Gamma_D^2 d\mu^D_1 \geq \int \Gamma_D^2 d\mu^D_2 = E \left[ \gamma_D^2 \right].$$

as desired.

Proof of Remark 1. Denote with $M^D_k$ player B’s initially expected monetary payoff in treatment $D_k$, i.e.,

$$M^D_k = \mathbb{E}_{\beta^D_k, \gamma^D_k} \left[ m^D_k \right] = \beta^D_k m^D_k + (1 - \beta^D_k) \gamma^D_k \ell.$$

We first show that $M^D_2 \geq M^D_1$ almost everywhere. Since $M^D_k (\beta^D_k, \gamma^D_k)$ is strictly increasing in $\gamma^D_k$, the following relation

$$M^D_2 (\beta^D_2, \gamma^D_2) = \beta^D_2 m^D_2 + (1 - \beta^D_2) \gamma^D_2 \ell \geq \beta^D_1 m^D_1 + (1 - \beta^D_1) \gamma^D_1 \ell = M^D_1 (\beta^D_1, \gamma^D_1)$$

holds for all $\gamma^D_1$ and $\gamma^D_2$ if and only if it holds with $\gamma^D_1 = 1$ and $\gamma^D_2 = 0$, i.e., if and only if

$$\beta^D_2 m^D_2 \geq \beta^D_1 m^D_1 + (1 - \beta^D_1) \ell.$$

Since for all $\beta^D_1$ and $\beta^D_2$ it holds that $\beta^D_2 \geq \beta^D_1 \geq \frac{1}{2} \geq (1 - \beta^D_1)$ and $(m^D_2 - m^D_1) >
\( \ell \), we obtain

\[
\beta^{D2} m_{b}^{D2} - \beta^{D1} m_{b}^{D1} \geq \beta^{D1} m_{b}^{D2} - \beta^{D1} m_{b}^{D1} = \beta^{D1} (m_{b}^{D2} - m_{b}^{D1}) \\
\geq \frac{1}{2} (m_{b}^{D2} - m_{b}^{D1}) \geq \frac{1}{2} \ell \geq (1 - \beta^{D1}) \ell,
\]

as desired.

With this, we now show that the frequency of acceptance of the greedy offer in treatment D1 is higher than in treatment D2. Player B with personal features \((\beta, \gamma, \theta, \delta)\) accepts the greedy offer in treatment Dk if and only if

\[
M_{b}^{Dk}(\beta, \gamma) \leq \hat{m}(\theta, \delta) := \ell + \theta h - \delta (h - \ell) \frac{\theta h}{\theta h}.
\]

Thus, we can compute the frequency of accepting the greedy offer in treatment Dk as:

\[
P(a^{Dk} = y) = \int \mathbf{1}_{M_{b}^{Dk}(\beta, \gamma) \leq \hat{m}(\theta, \delta)} \mu^{Dk}(d\beta, d\gamma, d\theta, d\delta)
\]

where \(\mu^{Dk} \in \Delta ([0, 1]^2 \times \mathbb{R}^2)\) is the treatment-dependent measure on all the personal features of B, including plan \(\gamma\).

To obtain

\[
P(a^{D1} = y) \geq P(a^{D2} = y),
\]

we first have to show that

\[
P(a^{D1} = y | \theta, \delta) \geq P(a^{D2} = y | \theta, \delta), \text{ i.e., that}
\]

\[
\int \mathbf{1}_{M_{b}^{D1}(\beta, \gamma) \leq \hat{m}(\theta, \delta)} \mu^{D1}(d\beta, d\gamma | \theta, \delta) \geq \int \mathbf{1}_{M_{b}^{D2}(\beta, \gamma) \leq \hat{m}(\theta, \delta)} \mu^{D2}(d\beta, d\gamma | \theta, \delta).
\]

Indeed, since \(M_{b}^{D2} \geq M_{b}^{D1}\) almost everywhere, \(\beta^{D2} \geq \beta^{D1}\), and \(M_{b}^{Dk}\) is decreasing in \(\beta\), it follows that

\[
\int \mathbf{1}_{M_{b}^{D1}(\beta, \gamma) \leq \hat{m}(\theta, \delta)} \mu^{D1}(d\beta, d\gamma | \theta, \delta) \geq \int \mathbf{1}_{M_{b}^{D2}(\beta, \gamma) \leq \hat{m}(\theta, \delta)} \mu^{D1}(d\beta, d\gamma | \theta, \delta) \\
\geq \int \mathbf{1}_{M_{b}^{D2}(\beta, \gamma) \leq \hat{m}(\theta, \delta)} \mu^{D2}(d\beta, d\gamma | \theta, \delta).
\]

Therefore, taking into account that the marginal distribution of \((\theta, \delta)\) is the same in the two treatments because they are personal traits (hence exogenous), we obtain

\[
P(a^{D1} = y) \geq \int \left( \int \mathbf{1}_{M_{b}^{D1}(\beta, \gamma) \leq \hat{m}(\theta, \delta)} \mu^{D1}(d\beta, d\gamma | \theta, \delta) \right) \text{ marg}_{\theta, \delta} \mu^{D1}(d\theta, d\delta) \\
\geq \int \left( \int \mathbf{1}_{M_{b}^{D2}(\beta, \gamma) \leq \hat{m}(\theta, \delta)} \mu^{D2}(d\beta, d\gamma | \theta, \delta) \right) \text{ marg}_{\theta, \delta} \mu^{D2}(d\theta, d\delta) = P(a^{D2} = y).
\]

\[\blacksquare\]
Proof of Lemma 2. We have to show that
\[ \int \Gamma_{D1} d\phi_{D1} \geq \int \Gamma_{D2} d\phi_{D2}. \]

Since \( m^2_b > m^1_b \) implies that \( \Gamma_{D1}(\beta, \theta, \delta) \geq \Gamma_{D2}(\beta, \theta, \delta) \) for all \( (\beta, \theta, \delta) \), it follows that
\[ \int \Gamma_{D1} d\phi_{D1} \geq \int \Gamma_{D2} d\phi_{D1}. \]

Then,
\[
\int \Gamma_{D2} d\phi_{D1} = \int \left( \int \Gamma_{D2}(\beta, \theta, \delta) \phi_{D1}(d\beta|\theta, \delta) \right) \text{marg}_{\theta, \delta} \phi_{D1}(d\theta, d\delta)
= \int \left( \int \Gamma_{D2}(\beta, \theta, \delta) \phi_{D1}(d\beta|\theta, \delta) \right) \text{marg}_{\theta, \delta} \phi_{D2}(d\theta, d\delta)
\geq \int \left( \int \Gamma_{D2}(\beta, \theta, \delta) \phi_{D2}(d\beta|\theta, \delta) \right) \text{marg}_{\theta, \delta} \phi_{D2}(d\theta, d\delta) = \int \Gamma_{D2} d\phi_{D2},
\]
where the second equality holds because personal traits \( \theta \) and \( \delta \) are treatment-independent, and the inequality follows from Assumption 2 and the fact that \( \Gamma \) is decreasing in \( \beta \). ■

Proof of Proposition 3. By Lemma 2, \( \alpha_{D1} \geq \alpha_{D2} \). This implies that
\[ P(\alpha_{D1} \leq \hat{\alpha}) \leq P(\alpha_{D2} \leq \hat{\alpha}), \]
i.e., it is less likely that a randomly drawn A-subject from treatment D1 has first-order belief \( \alpha \leq \hat{\alpha} \) than a randomly drawn A-subject from treatment D2. Thus, due to incentive compatibility, it is more likely to draw an A-subject that chooses the default allocation in treatment D2 than in treatment D1. ■

Proof of Proposition 4. It is easily verified that
\[ E[\gamma_{Dk}] = \int \Gamma_{Dk} d\mu_{Dk} \leq \int \Gamma_{Sk} d\mu_{Dk}, \]
because \( \Gamma_{Dk}(\beta, \theta, \delta) \leq \Gamma_{Sk}(\beta, \theta, \delta) \) almost everywhere. To see this, notice that \( \Gamma_{Sk} = 0 \) implies \( \Gamma_{Dk} = 0 \) because
\[ \hat{\delta}_{Sk}(\beta, 0, \theta) = \frac{\ell}{h - \ell} \geq \frac{\ell - h\theta \max\{0, \beta m_b - \ell\}}{h - \ell} = \hat{\delta}_{Dk}(\beta, 0, \theta). \]
Note that B’s probability of acceptance in treatment Sk, \( \Gamma_{Sk} \), depends only on the
exogenous personal trait $\delta$, hence it is treatment-independent:

$$
\int \Gamma^{Sk} d\mu^{Dk} = \int \Gamma^{Sk} d\mu^{Sk}.
$$

Hence,

$$
\mathbb{E} [\gamma^{Dk}] \leq \int \Gamma^{Sk} d\mu^{Sk} = \mathbb{E} [\gamma^{Sk}].
$$
as desired.

**Proof of Proposition 5.** First, we show that, given the payoff treatment $k$, A’s expectation that B accepts the greedy offer is lower in treatment $Dk$ than in treatment $Sk$. Indeed, for $k = 1, 2$ it holds that

$$
\alpha^{Dk} = \int \Gamma^{Dk} d\phi^{Dk} \geq \int \Gamma^{Sk} d\phi^{Dk} = \int \Gamma^{Sk} d\phi^{Sk} = \alpha^{Sk}
$$

where the inequality holds because $\Gamma^{Dk}(\beta, \theta, \delta) \leq \Gamma^{Sk}(\beta, \theta, \delta)$ almost everywhere and the second equality holds because $\Gamma^{Sk}$ depends only on the distribution of $\delta$, which is exogenous, hence treatment-independent. Therefore, for each $k$ it holds that $\alpha^{Sk} \geq \alpha^{Dk}$ almost everywhere, which implies that

$$
\mathbb{P}(\alpha^{Sk} \leq \hat{\alpha}) \leq \mathbb{P}(\alpha^{Dk} \leq \hat{\alpha}),
$$
i.e., it is less likely that a randomly drawn A-subject from treatment $Sk$ has first-order belief $\alpha \leq \hat{\alpha}$ than a randomly drawn A-subject from treatment $Dk$. Thus, due to incentive compatibility, it is more likely to draw an A-subject that chooses the default allocation in treatment $Dk$ than in treatment $Sk$. ■
Appendix B

Results by gender

This appendix reports gender specific results. Among B-subjects we had 112 males and 64 females. In the strategy method we had 37 males and 29 females, while in the direct method, out of 72 B-subjects who received the greedy offer, 48 were males and 24 females. Among A-subjects we had 101 males and 75 females.

Figure 7 illustrates acceptance rates by method of play and payoff treatment across genders. As for females, the average frequency of acceptance is lower in the direct method than in the strategy method (MW: p-value 0.01, 53 observations); there is mild evidence that the frequency of acceptance is higher in D2 than in D1 (MW: p-value 0.11; 24 observations) but there is no difference across S2 and S1 (MW: p-value 0.19, 29 observations). Considering males, the average frequency of acceptance is slightly lower in the direct method than in the strategy method (MW: p-value 0.19, 85 observations) but it is similar across payoff treatments, for fixed method of play (MW in the direct method: p-value 0.49, 48 observations; MW in the strategy method: p-value 0.86, 37 observations).

Figure 7: B-subjects’ behavior, by gender

Table 7 reports separately for each gender the average marginal effects of the treatments dummies on the probability of accepting the greedy offer resulting from a probit estimate of the model described in equation (5). In particular, we consider specifications I and III of Table 4 in the main text. Examining the average marginal effect of Direct in both specifications, it results that the effect of the direct method that we find for the whole sample is similar across genders: for both females and males, playing the direct response method rather than the strategy method decreases the probability of acceptance. Interestingly, the payoff treatment effect in the direct
method differs across genders. Specifically, Payoff Increase has a positive effect on the probability of acceptance for females and a negative effect on that of males. Yet, such conditional effect is significant only for females. Instead, in the strategy method, Payoff Increase has a small, and not significant, marginal effect on the probability of acceptance both for females and males.

Table 7: B-subjects’ behavior by gender

<table>
<thead>
<tr>
<th></th>
<th>Females</th>
<th></th>
<th>Males</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I  II</td>
<td>I  II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct</td>
<td>-0.328***</td>
<td>-0.354***</td>
<td>-0.140</td>
<td>-0.232**</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.113)</td>
<td>(0.104)</td>
<td>(0.093)</td>
</tr>
<tr>
<td>Payoff Increase</td>
<td>0.238**</td>
<td>0.195*</td>
<td>-0.045</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.114)</td>
<td>(0.104)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>Direct × Payoff Increase</td>
<td>0.191</td>
<td>0.260</td>
<td>-0.127</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>(0.233)</td>
<td>(0.226)</td>
<td>(0.209)</td>
<td>(0.183)</td>
</tr>
<tr>
<td>Belief</td>
<td>-0.006**</td>
<td>-0.007***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belief × Direct</td>
<td>-0.002</td>
<td>-0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trait Anger</td>
<td>-0.003</td>
<td>-0.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.008</td>
<td>-0.070**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.028)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payoff Increase in direct method</td>
<td>0.343*</td>
<td>0.335*</td>
<td>-0.101</td>
<td>-0.054</td>
</tr>
<tr>
<td></td>
<td>(0.193)</td>
<td>(0.185)</td>
<td>(0.143)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>Payoff Increase in strategy method</td>
<td>0.152</td>
<td>0.075</td>
<td>0.027</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.136)</td>
<td>(0.152)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>Observations</td>
<td>53</td>
<td>53</td>
<td>85</td>
<td>85</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is a dummy that indicates whether \( i \) accepts the greedy offer; the table reports average marginal effects estimated from probit models; for the treatments interaction, we report the differential marginal effects of Payoff Increase when Direct changes from 0 to 1 (conditional marginal effects are reported in the lower part of the table); for the interaction of Belief with Direct, we report the differential marginal effect of Belief when Direct changes from 0 to 1; standard errors in parentheses; significance levels are: * \( p < 0.1 \); ** \( p < 0.05 \); *** \( p < 0.01 \).

Next, we compare the treatments effects on A-subjects’ behavior across gender. Figure 8 shows, for females and males separately, the average frequency of choosing a default allocation by payoff treatment and method of play. When splitting the sample, the average frequency of the default allocation does not differ significantly across methods of play neither for females (MW: p-value 0.20, 75 observation) nor for males (MW: p-value 0.18, 101 observation). Moreover, in the strategy method the frequency of default allocations does not differ across payoff treatments neither
for females (MW: p-value 0.27, 34 observation) nor for males (MW: p-value 0.91, 32 observations). Instead, Payoff Increase has a significant effect in the direct method for both categories, but in opposite directions for females and males. In particular, the average frequency of default allocations for females is higher in D1 than in D2 (MW: p-value 0.05, 41 observation), while for males it is higher in treatment D2 than in treatment D1 (MW: p-value 0.09, 69 observations).

Figure 8: A-subjects’ behavior, by gender

Thus, only when limiting the attention to the male sample, A-subjects’ react to the payoff treatment as expected. This is confirmed by a probit estimate of the probability of choosing the default allocation, conducted separately for males and females, whose results are reported in Table 8 below. We adopt alternative specifications I and III used above for the whole sample of A-subjects (see Table 6). We find that the average marginal effect of Payoff Increase on the females’ probability of choosing a default allocation is negative and significant, which is at odds with our predictions. Instead, the average marginal effect of Payoff Increase for males, conditioning on the direct response method, is positive and mildly significant, while, when conditioning on the strategy method, it is positive but not significant.
Table 8: A-subjects’ behavior by gender

<table>
<thead>
<tr>
<th></th>
<th>Females</th>
<th></th>
<th>Males</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Direct</td>
<td>0.142</td>
<td>0.063</td>
<td>-0.127</td>
<td>-0.139</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td>(0.110)</td>
<td>(0.097)</td>
<td>(0.089)</td>
</tr>
<tr>
<td>Payoff Increase</td>
<td>-0.253**</td>
<td>-0.228**</td>
<td>0.124</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td>(0.105)</td>
<td>(0.084)</td>
<td>(0.077)</td>
</tr>
<tr>
<td>Direct × Payoff Increase</td>
<td>-0.119</td>
<td>-0.029</td>
<td>0.153</td>
<td>0.158</td>
</tr>
<tr>
<td></td>
<td>(0.223)</td>
<td>(0.215)</td>
<td>(0.193)</td>
<td>(0.175)</td>
</tr>
<tr>
<td>Belief</td>
<td>-0.006***</td>
<td>-0.007***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morality</td>
<td>0.004</td>
<td>0.024</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.072)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guilt aversion</td>
<td>0.031</td>
<td>0.013</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.034)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.047</td>
<td>-0.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.023)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payoff Increase in Direct</td>
<td>0.307**</td>
<td>-0.244*</td>
<td>0.172*</td>
<td>0.144*</td>
</tr>
<tr>
<td></td>
<td>(0.150)</td>
<td>(0.146)</td>
<td>(0.095)</td>
<td>(0.087)</td>
</tr>
<tr>
<td>Payoff Increase in Strategy</td>
<td>-0.188</td>
<td>-0.215</td>
<td>0.020</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.165)</td>
<td>(0.156)</td>
<td>(0.168)</td>
<td>(0.153)</td>
</tr>
<tr>
<td>Observations</td>
<td>75</td>
<td>75</td>
<td>101</td>
<td>101</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is a dummy that indicates whether i chooses the default allocation; the table reports average marginal effects estimated from probit models; for the treatments interaction, we report the differential marginal effects of Payoff Increase when Direct changes from 0 to 1 (conditional marginal effects are reported in the lower part of the table); standard errors in parentheses; significance levels are: * p < 0.1; ** p < 0.05; *** p < 0.01.

Appendix C

Experimental instructions: The following instructions are for all treatments, unless otherwise specified.

Welcome and thank you for participating in this experiment. By closely following the instructions you will have the chance to earn an amount of money that will be paid in cash at the end of the experiment. You are not allowed to talk or communicate with other participants. If you have any questions, raise your hand and an assistant will help you.

Instructions

The experiment relates to the two-player game described below. At the beginning of the experiment you will be randomly assigned the role of Player 1 or of Player
2. You will be paired randomly and anonymously with another participant with a different role. There will be 11 pairs of participants in total.

The game

Player 1 can make Offer A or Offer B to Player 2. Every offer is a way to assign an amount of money to each player. If Player 1 makes Offer A, the two players obtain automatically the amounts of money that this offer assigns to them. If Player 1 makes Offer B, Player 2 can either Accept it or Reject it.

If Player 1 makes Offer A

- Player 1 obtains 8 EURO.
- Player 2 obtains 8 EURO (D1 and S1). [Player 2 obtains 11 EURO (D2 and S2).]

If Player 1 makes Offer B and Player 2 Accepts it

- Player 1 obtains 14 EURO.
- Player 2 obtains 2 EURO.

If Player 2 Rejects it, both players get 0 EURO. The players’ choices jointly determine the outcome of the game and the players’ earnings as illustrated in the following Figure. Note that the first number in parenthesis indicates Player 1’s earnings, while the second indicates Player 2’s earnings.

(a) Figure in the instructions for D1 and S1

(b) Figure in the instructions for D2 and S2

How to make a choice (D1 and D2)

Player 1 decides first and chooses whether to make Offer A or Offer B, then his/her choice is communicated to Player 2.

Once informed about Player 1’s choice, Player 2 can find himself in either one of the following situations:
• if Player 1 made Offer A, Player 2 does not have to make any choice and the players obtain the amounts of money assigned to them by Offer A.

• if Player 2 made Offer B, Player 2 has to decide whether to accept or reject it and his decision determines the players’ earnings as described above.

[How to make a choice (S1 and S2)]
Player 1 decides whether to make Offer A or Offer B. Player 2 has to make a choice without knowing whether Player 1 has made Offer A or Offer B. Hence, he/she has to decide whether he/she would Accept or Reject Offer B in case he/she received it. Once both players have decided, their choices are combined and

• if Player 1 made Offer A, Player 2’s choice is ignored and the players obtain the amounts of money assigned to them by Offer A.

• if Player 2 made Offer B, Player 2’s choice applies and determines the players’ earnings as described above.]

You will now answer some control questions that will appear on the screen to verify whether you understood the instructions so far.

The computer has formed the 11 pairs and informed you about your role in the pair. With this role you will make a **CHOICE** in the game.

Before making a choice, you will be asked to make a **PREDICTION** about the choices that will be made in the game by participants of the **other 10 pairs** playing in a role different from yours.

Through your prediction you will have the chance to earn money as explained below. Note that you will be paid either what you earned in the prediction task or in the game. At the end of the experiment the computer will randomly determined for each pair whether the two players will be paid their earnings from the prediction task (probability 50%) or from the game (probability 50%). Earnings obtained in the selected part plus 4 EURO for filling a questionnaire will be paid in cash to you at the end of the experiment.

As the prediction tasks are different for the two players, we describe them separately.

**How to make a PREDICTION: Player 1 (D1 and D2)**
Consider participants in the role of Player 2 who have not been paired to you and have received Offer B. You will be asked to predict in which percentage they will Accept Offer B. You can express your prediction by indicating a percentage between 0% and 100% (rounding up to integer numbers). After all participants have made their choice in the game, you will be informed about the actual percentage of participants of the other pairs who accepted Offer B among those who have received it.

If the distance between the percentage you have predicted \(x\%\) and the actual one \(y\%\) is smaller than 10%, i.e., if

\[|x\% - y\%| < 10\%

you obtain 10 EURO, otherwise you obtain 2 EURO.

Note that you receive 2 EURO both in case you underestimate and in case you overestimate the actual percentage with a distance larger than 10 percentage points. Moreover, you obtain 2 EURO even if the distance is exactly equal to 10 percentage points.

In case none of the 10 participants have received Offer B, the computer will determine randomly (with probability 50%) whether you obtain 10 or 2 EURO.

[How to make a prediction: Player 1 (S1 and S2)

Consider the participants in the role of Player 2 who have not been paired to you. You will be asked to predict how many of them will Accept Offer B in case they received it.

You can express your prediction by indicating a number between 0 and 10.

After all participants have made their choice in the game, you will be informed about the actual number of participants of the other pairs who have decided that they would Accept Offer B in case they received it. If the number you indicated coincides with the actual number you obtain 10 EURO, otherwise you obtain 2 EURO.]

How to make a PREDICTION: Player 2

Consider the participants in the role of Player 1 who have not been paired to you. You will be asked to predict in which percentage they will make Offer A.

You can express your prediction by indicating a number between 0 and 10.

After all participants have made their choice in the game, you will be informed about how many participants of the other pairs made Offer A. If the number you indicated coincides with the actual number you obtain 10 EURO, otherwise you
obtain 2 EURO.

Summing up the timing of the experiment is as follows:

- Predictions;
- Choices in the game;
- You are informed about the earnings (both yours and of the other player) from the predictions and from the game;
- The computer selects either the game or the predictions for the payment and informs you about your final earnings;
- Questionnaire;
- Payment.

You will now answer some control questions that will appear on the screen to verify whether you understood the instructions.
Acknowledgements

We would like to thank Giuseppe Attanasi, Federico Bobbio, Lorenzo Cappellari, Anna Conte, Roberto Corrao, Martin Dufwenberg, Antonio Filippin, Werner Guth, Vittoria Levati, Emanuele Lo Gerfo, Paola Moscariello, Simona Sacchi, Dov Samet, Alec Smith and seminar participants at the Second Workshop on Psychological Game Theory held at the University of East Anglia and at the BEEN Workshop held at the University of Bologna for useful comments. Pierpaolo Battigalli gratefully acknowledges financial support from Bocconi University and ERC grant 324219.

References


and emotional responses in the Ultimatum Game task. *Cognition* 114:89-95.


