SFB 649 Discussion Paper 2014-057

A Tale of Two Tails: Preferences of neutral third-parties in three-player ultimatum games

Ciril Bosch-Rosa*

*Technische Universität Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

A Tale of Two Tails: Preferences of neutral third-parties in three-player ultimatum games

Ciril Bosch-Rosa*

First version: 13/2/2011 This Version: 31/10/2013

Abstract

We present a three-player game in which a proposer makes a suggestion on how to split \$10 with a passive responder. The offer is accepted or rejected depending on the strategy profile of a neutral third-party whose payoffs are independent from his decisions. If the offer is accepted the split takes place as suggested, if rejected, then both proposer and receiver get \$0. Our results show a decision-maker whose main concern is to reduce the inequality between proposer and responder and who, in order to do so, is willing to reject both selfish and generous offers.This pattern of rejections is robust through a series of treatments which include changing the "flat-fee" payoff of the decision-maker, introducing a monetary cost for the decision-maker in case the offer ends up in a rejection, or letting a computer replace the proposer to randomly make the splitting suggestion between proposer and responder. Further, through these different treatments we are able to show that decision-makers ignore the intentions behind the proposers suggestions, as well as ignoring their own relative payoffs, two surprising results given the existing literature.

JEL: C92, D71, D63, D31

Keywords: Ultimatum game, experiment, fairness, third party

*Department of Economics, Technische Universität Berlin. Email: cirilbosch@gmail.com. I am greatly indebted to Johannes Müller-Trede for running some of the Barcelona experiments and for endless discussion over Skype. I would like to thank Robin Hogarth and Eldar Shafir for guidance in the initial stage of this research, and Daniel Friedman, Nagore Iriberri, Rosemarie Nagel and Ryan Oprea for guidance at different stages. I would also like to thank Pablo Brañas and Teresa Garcia for their invaluable comments and help. Thank you also to Gabriela Rubio for all the help and support. I am also grateful to Circe Bosch for help in making this a readable manuscript. Finally I'd like to acknowledge that discussants at ESA meetings in Copenhagen and Tucson as well as at Universidad de Granada's seminar were of great help. This project was partially funded by the Deutsche Forschungsgemeinschaft (DFG) through the SFB 649 "Economic Risk"

1 Introduction

> "How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it." The Theory of Moral Sentiments, Adam Smith (1759)

Decisions made by uninvolved third-parties are not only an essential part of our judicial system, but are also central in many other more mundane activities. From a Supreme Court justice deciding over the Bush vs. Gore 2000 election results, to a building superintendent determining what neighbor is right in a noise complaint, neutral third parties impact our daily lives at many different levels. In fact, some studies claim that neutral third-parties should be ever more present in school conflicts as it promotes social cohesion and reduces bullying (Cremin (2007); Turnuklu et al. (2009)). Yet, very little work has been done on how the preferences of neutral third-parties look like.

In an effort to help shine some light on this topic we introduce a new three-player ultimatum game. In it a proposer makes an offer on how to split \$10 with a responder who plays no role in the game. Meanwhile, and without knowing the suggestion made by the proposer, a neutral decision-maker fills in a strategy profile accepting or rejecting all the potential offers the proposer can make. If the offer is accepted, then the split takes place as suggested; if rejected, then both proposer and responder get \$0. The decision-maker is paid a "flat fee" independent of his choices.

We use an ultimatum game setup to study this topic because a bargaining game is a simple way of modeling a neutral third-party intervention, and because we can use previous references as benchmark. In addition, in our ultimatum setup a rejection by the decision-maker leaves both proposer and responder with a \$0 payoff, which constitutes a strong disagreement signal on behalf of the neutral third-party.

The first result of our experiment shows that neutral decision-makers not only reject selfish offers, but they also refuse a substantial number of generous ones[1]. This appears to contradict previous results on three player games (Fehr and Fischbacher (2004) or Falk et al. (2008)) where neutral decision-makers rewarded generous offers and only punished selfish ones. To further look into it we introduce a series of robustness tests which include imposing a cost on the decision-maker if the game ends in a rejection, or having a computer replace the proposer, so that the splitting proposal (between proposer and responder) becomes random. These additional tests help us confirm our initial finding, and, most importantly, they show that proposers ignore the intentions behind a proposal, be them

[1] From now on we will consider any offer of more than \$5 to be "generous".

generous or selfish. It seems thus that decision-makers seem to care only for equality, making our results even more at odds with Fehr and Fischbacher (2004) or Falk et al. (2008).

If neutral decision-makers do ignore intentions as in our experiment, this should be of some concern for institutions that rely on neutral referees, as intentions of defendants play an important role in most legal systems (e.g. *mens rea* in criminal law, Martin (2003)). For instance, intentions are crucial in distinguishing between murder and manslaughter[2], and in most universities not only is cheating a violation of the honor code, but so is attempting to cheat. Whenever neutral decision-makers do not care about intentions and are only concerned about the final result of the game[3] it may be necessary to introduce some mechanism in our institutions to help correct the indifference towards the intentions of other players.

The paper is organized in the following way. In section 2 we cover the existing literature on the subject. Section 3 describes both the baseline game and the different treatments. Sections 4 describes our results. Section 5 discusses some methodological points of the experiment, and finally we conclude in section 6.

2 Literature Review

Three-player games are an essential part of the ultimatum game literature, and have been responsible for some key insights in the topic. In Knez and Camerer (1995), a proposer makes a simultaneous offer to two independent responders who can accept or reject proposals conditional on the offer made to the other receiver. The results show that responders are not willing to get offered less than their counterpart. In Güth and van Damme (1998), a proposer splits the pie with a decision-maker and a passive "dummy" player who plays no role in the game; if the offer is accepted, then the split goes as suggested, if rejected, then everyone receives zero. The result is that both proposer and responder end up ignoring the presence of the dummy player and split the pie between themselves. Finally, Kagel and Wolfe (2001) present us with a setup identical to Güth and van Damme (1998) except that now, if the offer is rejected, the dummy player gets a consolation prize. As in Güth and van Damme (1998), the dummy seems to play no role in the decision-makers mind, even when he gets a high consolation prize.

Many papers deal with the reasons behind the rejections of offers in two-player (and sometime three-player) bargaining games; from inequality aversion (Ockenfels and Bolton (2000) or Fehr and Schmidt (1999)), to punishment of selfish intentions (Blount (1995); Falk et al. (2005)), or Rawlsian preferences (Charness and Rabin (2002); Engelmann and Strobel (2004))[4], and even to the need for signaling discomformity (Xiao and Houser (2005)). But

[2] A distinction as old as 624 BCE when Draco drafted the first Athenian constitution and for the first time distinguished between these two terms Ehrenberg (1973)

[3] In our previous examples, whether or not someone is dead, or if the student actually copied or not.

[4] Which cannot really explain rejections in ultimatum games.

the literature grows silent when we look at the preferences of neutral third parties. Fehr and Fischbacher (2004) design a variation of the dictator game where a proposer offers an amount to a receiver, while a neutral third party can impose a (costly) punishment on the dictator. The results show that third party punishment is aimed to punish norm violators (i.e. selfish dictators) and not necessarily based on payoff differences among players.

On the other hand Leibbrandt and Lopez-Perez (2008) use a within-subject analysis which shows that second and third party punishment are driven by payoff differences rather than the intentions of the proposer. Interestingly, and against Fehr and Fischbacher (2004), they also find that second and third-party punishments are not significantly different in intensity[5]. More recently, Falk et al. (2008) have revisited the subject suggesting that while inequality has some effect on punishment, intentions of the proposer are the main reason behind most punitive actions. Our conclusions are in stark contrast with these latter results as we find that not only a significant number of generous offers are rejected by third parties, but that (against Blount (1995)) there are no statistical differences between the rejections to offers made by another subject, and those made randomly by a computer.

And, while we are not the first to report rejections of generous offers, we are the first to do so in a lab experiment. All previous reports of it were field experiments with subjects either from rural old Soviet Union regions (Bahry and Wilson (2006)) or small-scale societies in New Guinea (Henrich et al. (2001)). Furthermore, these previous results had always been 2 player games, and considered an anomalies. For example, Bahry and Wilson (2006) dismiss rejections of generous offers as a result of Soviet education, while Henrich et al. (2001) hypothesize that these rejections could be the result of a gift-giving culture, in which accepting large gifts establishes the receiver as a subordinate. Güth et al. (2007) also mention an inverted-U in ultimatum game data gathered through newspaper publications. Yet, they only informally mentioned it because of the small number of observations following this pattern.

Finally, there has been some controversy about the validity of the strategy method, a technique which we use in our experiment. Brandts and Charness (2011) is a good survey on the subject and supports the use of the strategy method. In fact, if we had used a direct method instead of the strategy method, the inverted-U results might have been even more prominent as Brandts and Charness (2011) report that punishment rates are lower if the strategy method is used. Further, Brandts and Charness (2011) claim that "in no case do we find that a treatment effect found with the strategy method is not observed in the direct-response method". See also Brandts and Charness (2000) for more information on the matter.

[5] Fehr and Fischbacher (2004) show that second-parties spend much more of their income to punish unfair dictators

3 Experimental Design

The experiment was run with a total of 282 undergraduates from both the Universitat Pompeu Fabra (UPF) in Barcelona, and the University of California Santa Cruz (UCSC) in Santa Cruz. Each session had 3 rounds and lasted on average 30 minutes. The mean earnings at UCSC were of \$4.5 and at UPF of €4.35 plus a show-up fee (\$5 and €3[6]) that was announced only at the end of the experiment[7]. Subjects were recruited through the ORSEE systems of each university (Greiner (2004)), and were required not to have any previous experience in bargaining games. In total 17 sessions were run, UCSC sessions had 12 subjects[8] and UPF sessions 18 subjects[9].

As subjects arrived to the lab, they were seated randomly in front of a terminal and the initial instructions were read aloud. In these instructions we announced that:

1. The experiment had three rounds and instructions for each round would be read *immediately before* each round started [10].
2. Each subject would be assigned a player type (A, B or C) which they would keep through the experiment.
3. Each round, subjects would be randomly assigned to a different group of three players (one of each type).
4. Only one of the rounds, randomly chosen by the computer, would be chosen for the final payoffs.
5. No feedback would be given until the end of the session[11], when they would be informed of the actions of subjects in their group for each round, as well as the round selected for the final payoffs.

Our experiment has a baseline treatment, and then 2 different robustness tests whose aim is to see how far we can push the results of the original treatment. Details on ordering and number of observations for each session can be found in Appendix A. A time line of the experiment is shown in Table 1.

[6]From now on, we will use the dollar sign to include both euros and dollars.

[7]While most subjects are aware of the rule of a "show-up fee" not announcing it until the end of the experiment adds pressure to the decision-makers would their decisions result in a rejection.

[8]Except 3 sessions that had 9 subjects.

[9]Except 2 sessions that had 12 subjects

[10]From experience, we prefer to read several times small amount of instructions rather than going over all instructions at the beginning of the session since subjects then get distracted. By breaking instructions into small concise parts we increase the likelihood that subjects are paying attention and, consequently, that they know what is expected of them in each round.

[11]This was done to minimize learning effects and have results of a "one-shot game" in each round.

Table 1: Steps of the experiment.

Step 1	Step 2	Step 3	Step 4
Read general instructions	Read instructions for Round 1	Round 1	Read instructions for Round 2
Assign player type	Assign players to group	No feedback	Assign players to new group

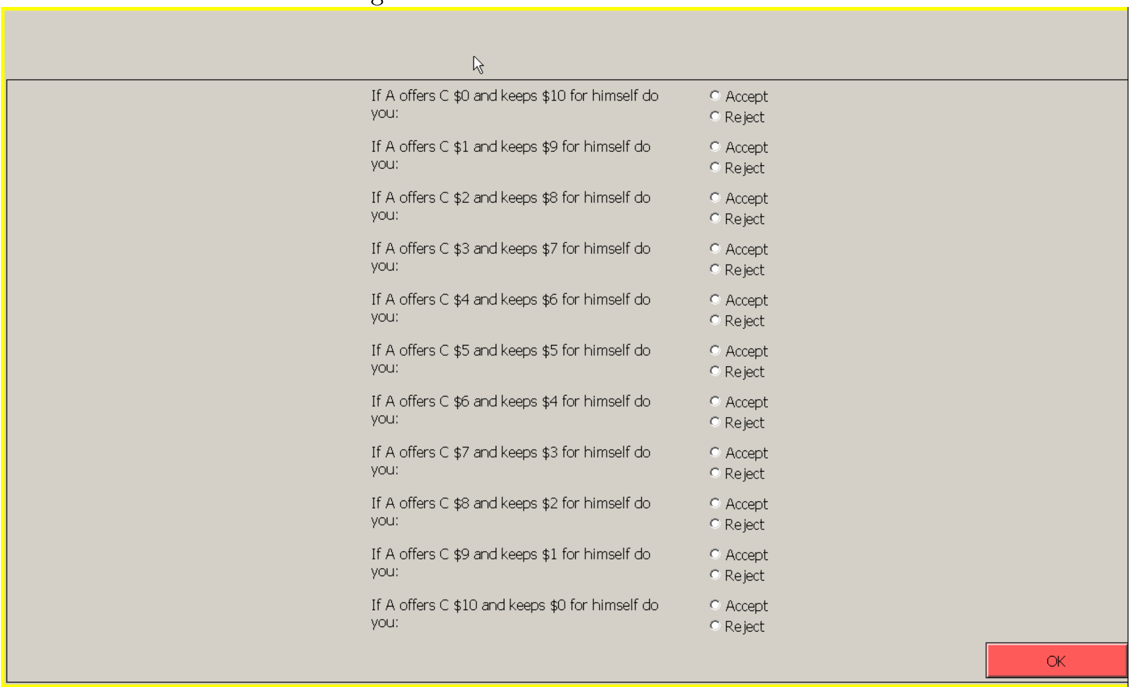
Step 5	Step 6	Step 7	Step 8
Round 2	Read instructions for Round 3	Round 3	Info on results for all games
No feedback	Assign players to group	No feedback	Final payoff info

3.1 Baseline

In the baseline design A players are assigned the role of proposer and have to make an offer on how to split \$10 with a C player who is a "bystanders" and plays no active role in the game. In the meantime (and without knowing the proposal made by A) B players are assigned the role of decision-makers and have to fill in a strategy profile accepting or rejecting all potential offers from A to C (screen-shot in Figure 1). If the offer is accepted, the split goes as suggested by A; if rejected, then both A and C get \$0 for the round. B player payoffs will be the treatment variable and these are:

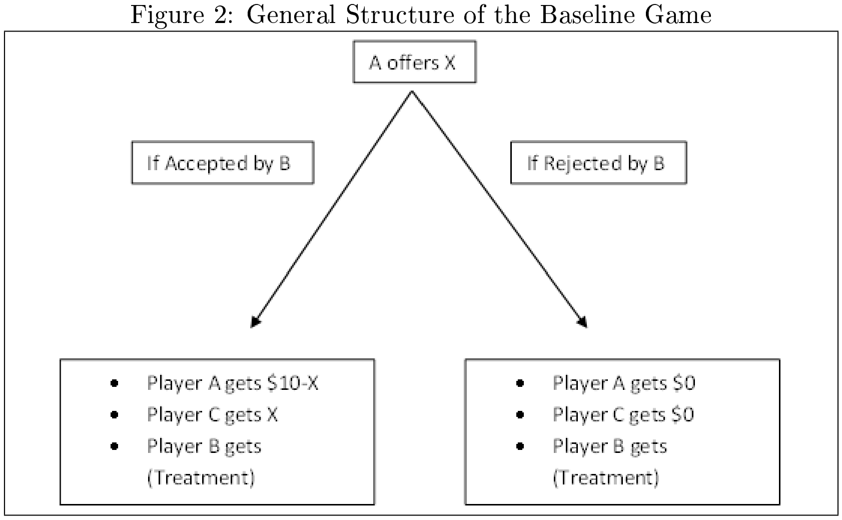
- Low (L): B gets paid \$3 for his decisions, whatever the outcome of the game
- Normal (N): B gets paid \$5 for his decisions, whatever the outcome of the game.
- High (H): B gets paid \$12 for his decisions, whatever the outcome of the game.

Figure 1: Decision-Maker Screen-shot



Treatments L, H, and N allow us to test whether or not decision-makers take into account their relative payoff when making the accept/reject decision. If no differences can be observed across treatments, then it will mean that we are observing the revealed preferences of a subject who has truly no strategic or monetary concerns in the game; what Fehr and Fischbacher (2004) call "truly normative standards of behavior". Figure 2 graphically lays out the general structure of the baseline game.

Figure 2: General Structure of the Baseline Game



3.2 Robustness tests

Our robustness tests are variations of the baseline and were introduced to put to a test the unexpected results of our original treatments. In order to do this, we will use the H and L treatments of the baseline and adapt them to our new games, while using N will be used as the measure to which we will compare all the different treatments in the experiment.

3.2.1 Costly rejection

In this robustness test we have the exact same setup as the baseline, except that now if the game ends in a rejection, then the decision-maker is penalized by a subtraction of $1 from his payoff for this round. So the treatments in the costly costly rejection sessions are:

- Low (L-1) : B gets paid $3 if A's offer is accepted and $2 if rejected.
- High (H-1) : B gets paid $12 if A's offer is accepted and $11 if rejected.

3.2.2 Computer

In the "computer" robustness test we have the same setup as in the baseline treatment, but this time the suggestion on how to split the $10 is randomly[12] made by the computer. This leaves both A and C as bystanders[13], while B fills in his strategy profile as usual. If the offer is accepted, then the split takes place as suggested by the machine, if rejected, then both A and C get $0. B's payoffs are completely independent from his choices and are:

- Low (Lm): B gets paid $3 for his decisions, whatever the outcome of the game.
- High (Hm): B gets paid $12 for his decisions, whatever the outcome of the game.

3.3 2UG

Finally, in all sessions one of the rounds will be the 2UG game. This game is designed to be a regular ultimatum game but keeping the 3-player group structure, as now A makes *two independent suggestions* on how to split $10;

[12] Following a uniform distribution across the whole offer space.

[13] Note that A is still a (human) subject getting a payoff that depends on the decisions made by B and the random split suggested by the computer.

one to B, the other to C. As in the baseline, we use the strategy method to elicit both B and C's preferences over the offers made *to them*. If B (C) rejects the offer that A made *to him*, then B (C) gets $0 for the round. If, instead B (C) accepts the offer, then the split goes as suggested by A. A's payoff is randomly chosen from one of the two different outcomes; if the selected game turns out to be a rejection, then A gets $0 for the round, if an acceptance, then A gets his part of the proposal. The purpose of randomizing A's payoffs is to prevent portfolio effects and to make payoffs fair across all subject types.

The 2UG game is introduced in our sessions for three reasons. The first one is to create a "break" between our treatments of interest[14] and so be able to recreate a "first-shot" scenario in the third round of the session. Secondly we use the 2UG as a control for our population sample, and to verify whether or not our subjects understand the strategy method interface. Finally, and very important for our results, the 2UG game shows that decision-makers take seriously the possibility of generous offers when filling out their strategy profile.

4 Results

The analysis of our data begins by looking at the baseline treatments in section 4.1, to then study the results of both robustness tests in section 4.2. Finally we discuss our general results and experimental design in section 5, and conclude in section 6. The 2UG outcomes can be found in Appendix B, where we show that our sample is not different from that of any other ultimatum-game experiment, and that subjects understand perfectly the instructions and interface.

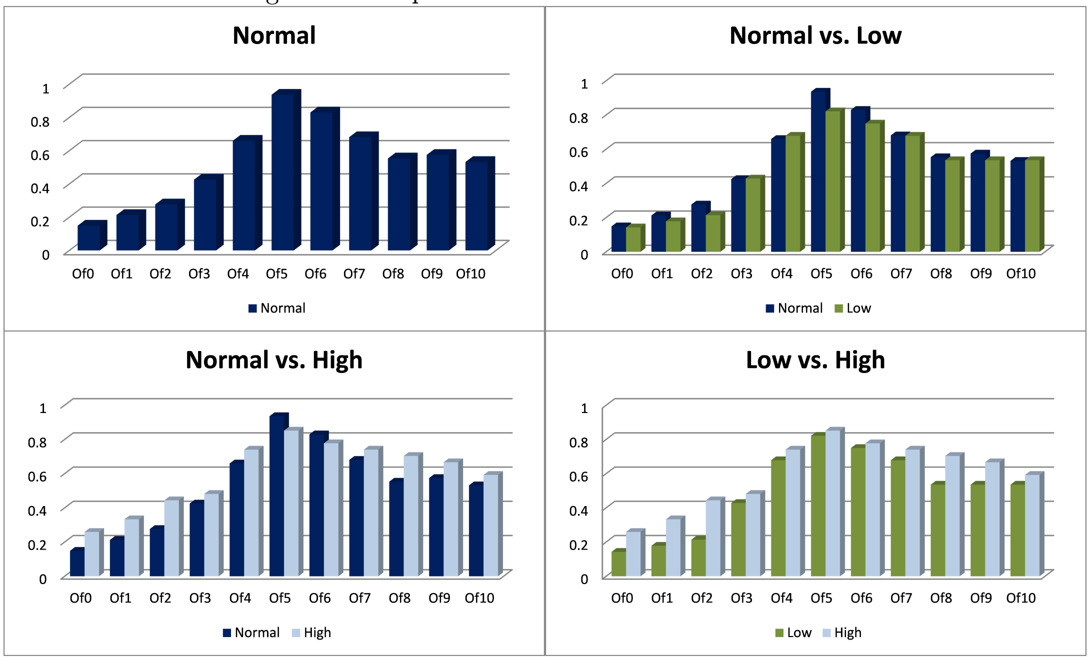
4.1 Baseline

Figure 3 presents the percentage of acceptances for each potential offer. In the upper-left corner we see the treatment N and in a clockwise order the comparison between N and L, L and H, and finally between N and H in the lower-left quadrant. Two things stand out immediately from these graphs. First, in all treatments there is a significant amount of rejections of both selfish *and* generous offers. In fact, if an offer is generous, the more generous it is, the *less* likely that it will be accepted[15].

Second, whether we pay a flat-fee of $3 or $12, both treatments show a very similar pattern of rejections. In fact, the rates of acceptance for each offer are not statistically different (Results for a Two-sided Fisher test can be found in Appendix C), and subjects seem to be consistent in their choices across treatments. A Wilcoxon matched-pairs signed-rank test presents us with no statistical differences when comparing the number of acceptances made by the same subject participating in an N or an L treatment (p-value = 0.375) nor among those taking part in N and H

[14]Some 2UG rounds are at the beginning of the session just to show that there are no ordering effects.
[15]Thus creating the inverted-U shape that Bahry and Wilson (2006) first identified in their field experiments.

Figure 3: Acceptance Rates for Baseline Treatments



(p-value = 0.161)[16], showing that decision-makers have stable preferences across treatments.

To further analyze our results, we run a regression of total accepted offers (Total) on dummies for location (Where), order (First), and treatment (High and Low). The results are in Table 2, with the first two columns comparing treatment H to N, and L to N respectively. In the third and fourth columns, H and L are compared together to the N treatment. The results show that payoffs and order of treatments have no effect on the number of accepted offers[17], and neither does location (all of these results are later confirmed in Table 3).

Table 2: Regression of total accepted offers by subject and treatment.

	(1) Total	(2) Total	(3) Total	(4) Total
Low	-1.093	-1.327	-0.330	0.165
	(0.848)	(0.817)	(0.666)	(0.796)
First		1.707*		1.008
		(0.947)		(0.717)
Where		-0.101		-0.263
		(1.281)		(0.979)
High			0.763	1.399
			(0.805)	(0.978)
cons	6.593***	6.318***	5.830***	5.114***
	(0.747)	(0.637)	(0.461)	(0.816)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

It is thus apparent that decision-makers do not take their own payoff into account when making decisions, which

[16] On the other hand, the test becomes somewhat more significant when comparing L and H (p = 0.0825), probably because the number of subjects participating in both H and L is extremely low (n = 4). See Appendix D for a lengthier discussion on this question.

[17] Column 2 shows some minimal order effects. We attribute these to the lack of first round H treatment observations. See Appendix D.

means that with our game design we are able to study the preferences of a neutral third-party with no strategic or monetary concerns in the game.

- **Result 1:** *In the baseline game there is no statistical difference in rejection patterns across the different treatments, indicating that decision-makers ignore their payoffs when making decisions.*

To better understand the data we define "absolute inequality" as the absolute value of the difference between A and C's payoff. Then we label all offers to the left of $5 (the "selfish" offers) as those in the Left-Hand-Tail (LHT), and all offers to the right of $5 (the generous offers) as those in the Right-Hand-Tail (RHT). A Spearman rank correlation test (Appendix E) shows a strong positive (and monotonic) relationship between the increase in absolute inequality and the rejection rate, which means that in both tails, the bigger the inequality in the split, the lower the chance of the offer being accepted.

We also run a linear probability model[18] (Table 3) where the binary accept/reject outcome is the dependent variable, and we have dummies for order (First), treatment (High, Low), location (Where), as well as dummies for distance. The coding for the distance dummies includes the distance to the even split and the tail they are in. So, for example, dist3l is the dummy for the $2 offer (which is 3 dollars to the left of $5) and dist2r is the dummy for an offer of $7 (which is 2 dollars to the right of $5). Column 5 of Table 3 has the full specification of the regression, and as we can see that all dummies for distance are negative and highly significant. Moreover, if we look at the coefficients for the distance dummies, the further away an offer is from $5 the lower the probability of being accepted. This relationship is monotonic in both tails [19] ranging from an 8% lower probability of acceptance for an offer of $6 (dist1r) to a 33.3% lower probability of acceptance for an offer of $10 (dist5r) when comparing them to the probability of acceptance of the even split $5.

As we can see, the decision-maker's preferences for equality are so strong that not only are they willing to leave the proposer and responder with a $0 payoff when the offer is selfish, but they are also willing to leave them with $0 if the offer is too generous.

- **Result 2:** *The greater the absolute inequality the lower the probability of the proposal being accepted.*

However, in Figure 3 we see that the inverted-U is not perfectly symmetric around the fair split, as there is a higher number of acceptances in the RHT (generous offers) than in the LHT (selfish offers). This might mean

[18] With clustered errors at the individual level.

[19] Strictly monotonic in the LHT and weakly in the RHT, confirming the Spearman rank correlation results.

Table 3: Linear Probability model of Accepted Offers.

	(1) Accept	(2) Accept	(3) Accept	(4) Accept	(5) Accept
Low	-0.0300	0.0150	0.0150	0.0150	0.0150
	(0.0547)	(0.0671)	(0.0673)	(0.0673)	(0.0674)
High	0.0693	0.127	0.127	0.127	0.127
	(0.0617)	(0.0803)	(0.0805)	(0.0805)	(0.0806)
First		0.917	0.917	0.917	0.917
		(0.0581)	(0.0582)	(0.0581)	(0.0584)
Where		-0.0239	-0.0239	-0.0239	-0.0239
		(0.0895)	(0.0897)	(0.0897)	(0.0899)
Dist1l			0.00327		-0.196
			(0.0501)		(0.0469)
Dist2l			-0.242***		-0.441***
			(0.0587)		(0.0625)
Dist3l			-0.379***		-0.578***
			(0.0558)		(0.0636)
Dist4l			-0.448***		-0.647***
			(0.0555)		(0.0603)
Dist5l			-0.507***		-0.706***
			(0.0578)		(0.0849)
Dist1r				0.340***	-0.088***
				(0.0447)	(0.0306)
Dist2r				0.242**	-0.186***
				(0.0493)	(0.0662)
Dist3r				0.134**	-0.294***
				(0.0530)	(0.0531)
Dist4r				0.134**	-0.294***
				(0.0536)	(0.0584)
Dist5r				0.0948*	-0.333***
				(0.0565)	(0.0625)
Cons	0.530***	0.465***	0.608***	0.379***	0.807***
	(0.0415)	(0.0739)	(0.0803)	(0.0708)	(0.0690)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

that decision-makers care about the intentions of proposers. To check the extent of this asymmetry, we run a linear probability model for each treatment (H, N, L), and compare the coefficients of the offers with same absolute inequality through a Wald Test (Table 4). The results show that in all the treatments the tails are asymmetric, with a higher degree or rejections on the selfish side (LHT). A Two-sided Fisher test comparing the number of accepted offers for same absolute inequality proposals confirms this result (Appendix F).

Table 4: P-values of Wald test for equality in within treatment regression coefficients.

Treatment	dist1l=dist1r	dist2l=dist2r	dist3l=dist3r	dist4l=dist4r	dist5l=dist5r
L	0.3357	0.0187***	0.0052***	0.0026***	0.0013***
H	0.5813	0.0066***	0.0186***	0.0016***	0.0016***
N	0.0107**	0.0021***	0.0022****	0.000***	0.000***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- **Result 3:** *In the baseline treatments, decision-makers are less willing to tolerate inequality when it is the result of a selfish offer.*

The three results presented above offer a picture of a decision-maker who does not seem to care about his relative payoff, but who is extremely concerned with the inequality between proposer and responder, as well as showing some dislike for selfish offers.

4.2 Robustness Tests

In this sections we analyze both robustness tests. The first one is the "costly-rejection" game. The design is identical to the baseline, except that now the decision-maker has to pay a $1 penalty if the game ends in a rejection[20]. This test was introduced to put downward pressure on the number rejections that B players make. If we still observe rejections of both selfish and generous offers in spite of the penalty, then this may be taken as a strong indication of the commitment of decision-makers towards equality. Also, the introduction of this penalty allows us to get a sense of what type of concerns, whether intentions of the proposer or absolute inequality, are more fragile in the decision-maker's preference set. If it happens that intentions play a stronger role than absolute inequality aversion, then we should observe acceptance in the RHT go up relative to those in the LHT. On the other hand, if inequality aversion is more important than intentions, then the result of introducing a penalty should be a much more symmetric pattern of rejections in this treatment than in the baseline game.

The second robustness check is the "computer" game. Again, we maintain the baseline design, but now the offer from A to C will be randomly chosen by a computer[21]. Because there are no intentions ingrained in the offers, but there still might be inequality, we expect to find a symmetric distribution of acceptances around the even split. Yet, what will be important in this game is the statistical comparison between the baseline and the computer treatment; if there is no statistical difference between them, then it will mean that intentions have very little weight in the decision-maker's preferences. If there is, then it will mean that intentions are (significantly) important for decision-makers.

4.2.1 Costly-Rejection

In Figure 4 we present the results of the costly-rejection treatments and compare them to their baseline counterparts and with the N treatment. The first thing that catches our attention is that, even when rejections are costly to decision-makers, we still observe them in both tails, following the same negative monotonic pattern that we already saw in the baseline treatment [22].

[20]Details can be found in section 3.2.1.
[21]Details can be found in section 3.2.2.
[22] See Appendix G for Spearman Correlation results

Figure 4: Acceptance rates of L-1 and H-1 vs. Normal, Low, and High Treatments

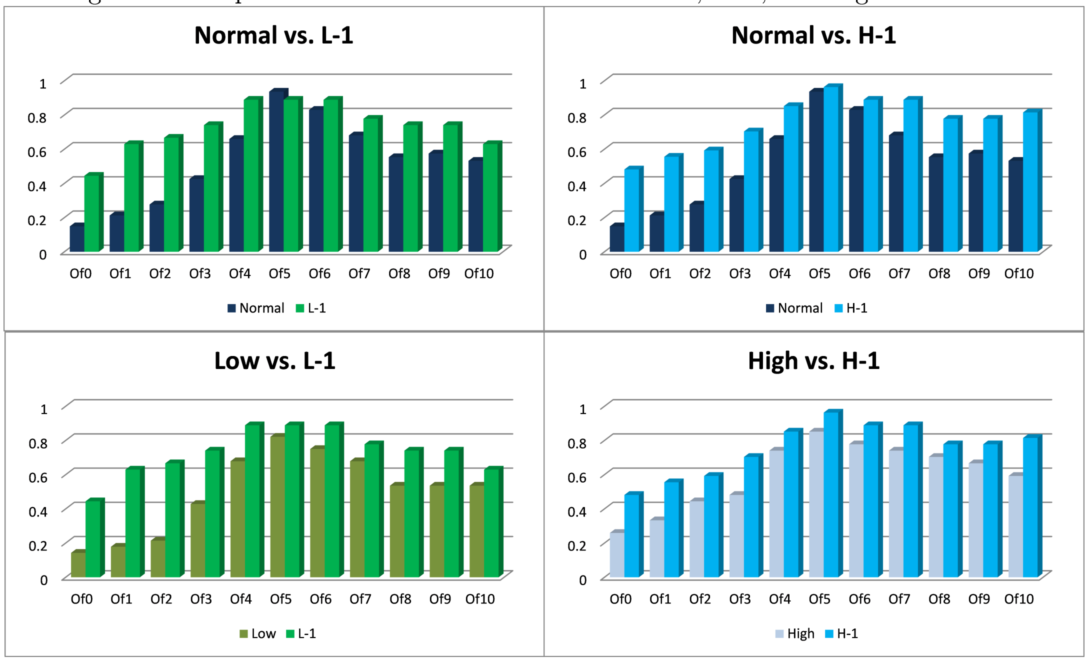
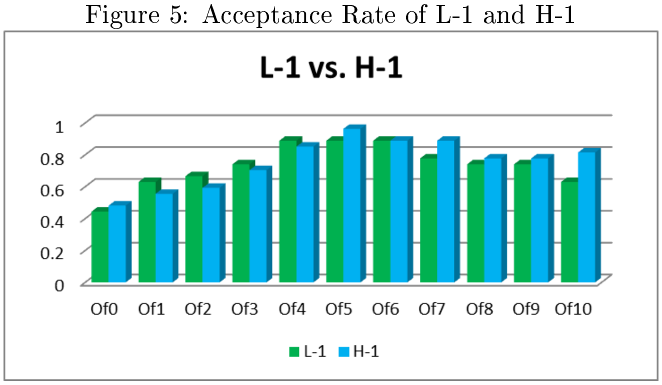


Figure 5: Acceptance Rate of L-1 and H-1



Furthermore, the similarity between H-1 and L-1 is striking (detail in Figure 5). Running a Wilcoxon matched-pairs sign-rank test comparing the number of accepted offers in each treatment, we find that the decision-maker's behavior is not statistically different across treatments (p-value = 0.6172). Additionally, both the linear probability model of Table 5 and a Two-sided Fisher test (Appendix H), confirm that there exists no significant difference between treatments. So, even when the relative costs of rejecting offers are wide apart, decision-makers behave in a similar manner under both costly treatments.

- **Result 4:** *Even with widely different relative rejection costs, there is no significant difference across treatments in the Costly-Rejection game.*

Table 5: Linear Probability model of Accepted Offers.

	(1) Accept	(2) Accept	(3) Accept	(4) Accept	(5) Accept
High1	0.0236	0.0289	0.0289	0.0289	0.02889
	(0.0615)	(0.0671)	(0.0674)	(0.0674)	(0.0677)
First	0.0236	0.0289	0.0289	0.0289	0.02889
	(0.0615)	(0.0671)	(0.0674)	(0.0674)	(0.0677)
Dist1l			0.0556		-0.0556
			(0.0417)		(0.0412)
Dist2l			-0.0926***		-0.204***
			(0.0391)		(0.0675)
Dist3l			-0.185***		-0.296***
			(0.643)		(0.0820)
Dist4l			-0.222***		-0.333***
			(0.0595)		(0.0809)
Dist5l			-0.352***		-0.463***
			(0.0697)		(0.0849)
Dist1r				0.188***	-0.0370
				(0.0574)	(0.0461)
Dist2r				0.133**	-0.0926
				(0.0493)	(0.0662)
Dist3r				0.0585	-0.167**
				(0.0500)	(0.0763)
Dist4r				0.0585	-0.167**
				(0.0462)	(0.0660)
Dist5r				0.0216	-0.204***
				(0.0487)	(0.0675)
Cons	0.731***	0.713***	0.786***	0.672***	0.897***
	(0.0608)	(0.0840)	(0.0809)	(0.0893)	(0.0860)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

On the other hand, we do see some differences when comparing the costly rejections treatments and their baseline counterparts. Running a regression on total accepted offers comparing H to H-1 and L to L-1 we see significant differences (p= 0.002 and p = 0.000 respectively) for their treatment dummies.

From Figure 4 it seems that most differences across baseline and robustness treatments stem from an increase of acceptances in the LHT. Apparently, when a cost is introduced, decision-makers accept relatively more selfish offers, while keeping a similar rate of rejections for the generous ones. To test this interpretation, we run a one-sided Fisher test comparing the number of accepted offers for each potential splitting suggestion, and confirm that the differences are mostly in the LHT (Table 6). Therefore we conclude that decision-makers have only some weak concern for the intentions of the proposer, while their absolute inequality aversion seems pretty robust, as the introduction of a cost has almost no effect on the latter but it increases significantly acceptance rates in the former.

This conclusion is supported by a Wald test comparing the coefficients of the offers that have the same level of absolute inequality (Table 7). The results show a symmetric L-1, but a slightly unbalanced H-1. A Two-sided Fisher (Table 8), shows symmetry under both treatments.

Table 6: One-sided Fisher P-values comparing total acceptances per treatment.

	\$0	\$1	\$2	\$3	\$4	\$5	\$6	\$7	\$8	\$9	\$10
L vs. L-1	0.01**	0.01**	0.01**	0.01**	0.05*	0.37	0.16	0.30	0.09*	0.09*	0.33
H vs. H-1	0.07*	0.08*	0.20	0.08*	0.25	0.17	0.23	0.14	0.37	0.27	0.06*

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: P-values of Wald test

Treatment	dist1l=dist1r	dist2l=dist2r	dist3l=dist3r	dist4l=dist4r	dist5l=dist5r
L-1	1.00	0.7536	0.5302	0.3466	0.1175
H-1	0.7410	0.0991*	0.0991*	0.0481**	0.0032***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- **Result 6:** *Under Costly-Rejection treatments, the intentions of the proposer play a minor role in the acceptance pattern of decision-makers, its impact disappearing completely in the costlier case (L-1).*

4.2.2 Computer Treatment

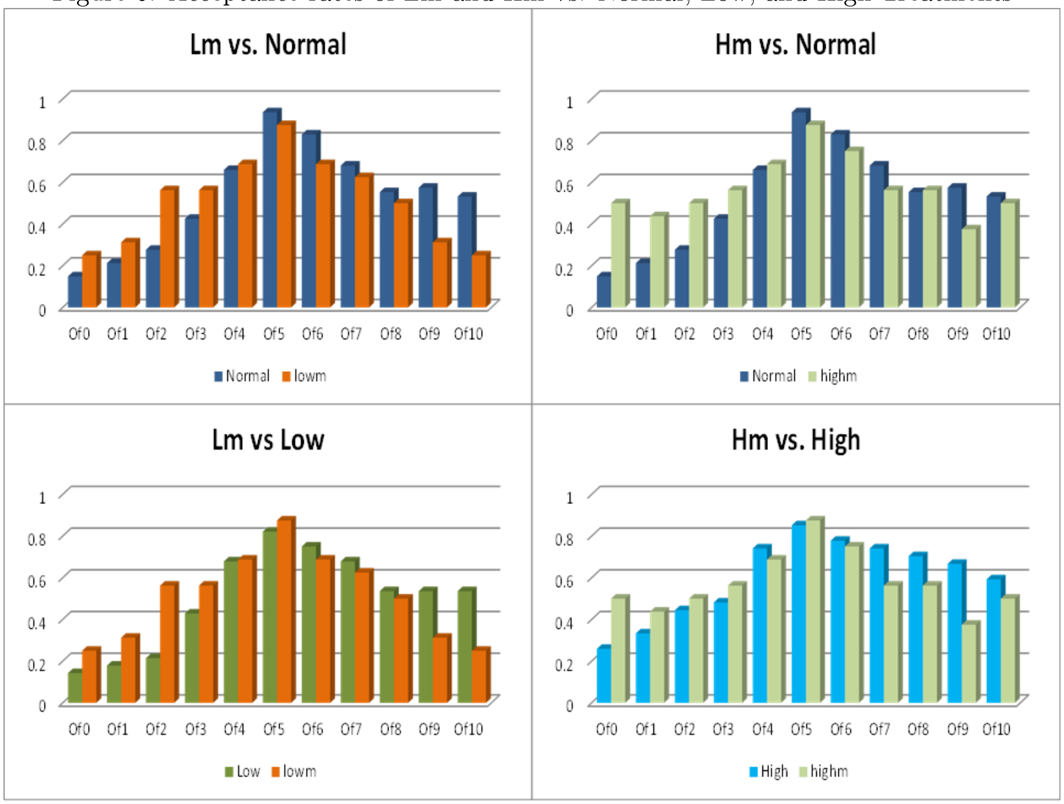
Figure 6 presents the results of this game and compares Hm to N and H in the right column, and Lm to N and L in the left column. It is apparent that all computer treatments are symmetric around the fair split (a two-sided Fisher test comparing equal absolute inequality offers shows a p-value=1.000 for all cases in both treatments), so decision-makers do not differentiate between the RHT and the LHT in this game.

Table 8: Two-Sided Fisher P-values.

Treatment	\$4=\$6	\$3=\$7	\$2=\$8	\$1=\$9	\$0=\$10
L-1	1.000	1.000	0.766	0.559	0.275
H-1	1.000	0.175	0.241	0.148	0.021**

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 6: Acceptance rates of Lm and Hm vs. Normal, Low, and High Treatments



To compare the machine and baseline treatments we run a regression of total number of accepted offers on a dummy for treatment, location, and ordering in Table 9. The first three columns compare the data of Lm to Hm, then to L, and finally to N. The fourth and fifth column compare Hm first to H and then to N, and the last column compares Hm and Lm together to the N treatment results. There appears to be no significant differences across treatment dummies, so independent of whether the offer was made by a human or a computer, the number of accepted offers are the same. A two-sided Fisher test in Appendix I confirms this result, as does column 3 of Table 9, where we run a linear probability model comparing the data of both Lm and Hm to N.

- **Result** 7: *Rejection patterns of offers made (randomly) by a computer are not statistically different from rejection patterns of offers made by a human being.*

Table 9: Regression of total accepted offers by subject and treatment.

	(1) Total	(2) Total	(3) Total	(4) Total	(5) Total	(6) Total
Lm	-0.320	-0.0286	-0.527			-0.395
	(0.701)	(1.542)	(1.488)			(1.367)
Hm				-1.182	0.155	0.206
				(2.326)	(1.370)	(1.373)
First	0.110	0.357	-0.717	1.918	-0.724	-0.530
	(0.594)	(1.098)	(0.956)	(1.593)	(1.105)	(0.648)
Where		0.499	0.393	-0.554	0.398	0.274
		(1.366)	(1.206)	(1.679)	(1.256)	(1.126)
cons	6.053***	5.189***	6.109***	6.463***	6.112***	6.041***
	(1.356)	(0.774)	(0.653)	(0.821)	(0.684)	(0.598)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

So, while Result 3 found that intentions are somewhat important to decision-makers, Result 6 and Result 7 show that the weight that these have on acceptance rates is really small, especially when compared to how important absolute inequality appears to be.

To offer an overall picture of this experiment, Table 10 presents a linear probability model comparing N to the high and low payoff treatments of each game.[23] The results show a significant difference of the costly rejection treatments when compared to the N baseline (both Low1 and High1 are significantly different at the 5% in column 2), while neither the computer treatment nor the baseline treatments H and L are different from N. In all cases, again, distance from the fair split is highly significant and the probability of acceptance decreases monotonically as offers get away from the fair split in either direction.

In summary, after testing the preferences of decision-makers across three different games, it is pretty cleat that the main motivation for rejecting an offer is to reduce the payoff inequalities between players A and C. Intentions of the proposer, on the other hand, have only a minor effect in the baseline treatments where acceptance distributions are not perfectly symmetric around the $5 even split. Finally, when a cost for rejecting a proposal is introduced, rejection patterns differ from the pattern baseline treatment N, yet we do not observe significant differences between treatments H-1 and L-1.

[23]Please note that the p-value notation is changed in this table with respect to all other tables in the paper.

Table 10: Linear probability model comparing each treatment to baseline N treatment

	(1) Baseline	(2) Costly	(3) Computer
first	0.0917	0.0211	0.0164
	(0.0584)	(0.0581)	(0.0613)
where	-0.0239	-0.0121	-0.0107
	(0.0899)	(0.0956)	(0.0958)
low	0.0150		
	(0.0674)		
high	0.127		
	(0.0806)		
low1		0.213^{*}	
		(0.101)	
high1		0.241^{*}	
		(0.105)	
lowm			-0.00436
			(0.119)
highm			0.0540
			(0.122)
dist1r	-0.0882^{**}	-0.0693^{*}	-0.127^{*}
	(0.0306)	(0.0325)	(0.0545)
dist1l	-0.196^{***}	-0.158^{***}	-0.241^{***}
	(0.0469)	(0.0400)	(0.0648)
dist2r	-0.186^{***}	-0.168^{***}	-0.266^{***}
	(0.0467)	(0.0470)	(0.0667)
dist2l	-0.441^{***}	-0.347^{***}	-0.430^{***}
	(0.0625)	(0.0528)	(0.0735)
dist3r	-0.294^{***}	-0.267^{***}	-0.367^{***}
	(0.0531)	(0.0539)	(0.0663)
dist3l	-0.578^{***}	-0.465^{***}	-0.532^{***}
	(0.0636)	(0.0584)	(0.0761)
dist4r	-0.294^{***}	-0.257^{***}	-0.430^{***}
	(0.0584)	(0.0494)	(0.0612)
dist4l	-0.647^{***}	-0.515^{***}	-0.633^{***}
	(0.0603)	(0.0576)	(0.0616)
dist5r	-0.333^{***}	-0.297^{***}	-0.443^{***}
	(0.0625)	(0.0505)	(0.0651)
dist5l	-0.706^{***}	-0.614^{***}	-0.671^{***}
	(0.0584)	(0.0564)	(0.0635)
_cons	0.807^{***}	0.804^{***}	0.897^{***}
	(0.0690)	(0.0670)	(0.0686)
N	1122	1111	869
adj. R^2	0.193	0.180	0.142

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5 Discussion

Neutral referees make complex decisions based on a number of different factors. Fehr and Fischbacher (2004) and Falk et al. (2008) postulate that (selfish) intentions the punishment of proposers. In Leibbrandt and Lopez-Perez (2008), on the other hand, envy is identified as the main reason behind third-party punishment. In our experiment we find that the main concern of neutral third-parties is avoiding absolute inequality between proposer and responder. In fact, we find that this concern is so strong that decision-makers are willing to punish both proposer and responder with a $0 payoff if the offer is too generous, to avoid too big of an inequality between them.

Much of our data analysis is aimed at showing that there is no learning and no ordering effects in our results. This is necessary precaution because we collect the data following a within-subject design, yet we use them as if they came from a between-subject experiment. The reason is that in a between-subject design we would have collected only one observation for every three subjects invited into the lab, making the experiment expensive and time-consuming. Thankfully, having managed to show that there are no ordering or learning effects, we can use our data as if they all came from "first-shot" interactions. We believe that not giving feedback until the end of the session, mixing groups between rounds, paying only one round, and having the 2UG "break" are all crucial tools to avoiding any learning in our subjects.

Finally, we would want to mention that even though the number of observations for the computer treatment is not large, the results appear to be robust when tested in different ways.

6 Conclusion

Neutral third parties are everywhere in our institutions: from the members of the European Commission[24] deciding how to allocate the farming subsidies, to the referees of the Super Bowl, to the TV show "Judge Judy"[25]. Yet, as important as neutral third-parties are, the literature studying their preferences is still slim.

In an effort to help shine some light on this topic, we run an experiment introducing a new version of the three-player ultimatum game. In it a proposer makes an offer on how to split $10 with a responder who plays no role in the game. Meanwhile, and without knowing the suggestion made by the proposer, a neutral decision-maker fills in a strategy profile accepting or rejecting all the potential offers from the proposer. If the actual offer is accepted, then the split goes as suggested; if rejected, then both proposer and receiver get a payoff of $0. The payoff of the neutral third-party is independent of his decisions.

[24]Note that even if there is one commissioner per member state, these are expected to represent the interests of the EU and not their respective countries.

[25]This is a program where a retired judge decides over small-claim disputes, and where both plaintiff and defendant have previously signed a contract agreeing to accept the resolution of the "judge".

The results of the experiment show that neutral decision-makers are mostly concerned with reducing the payoff differences between proposer and responder, even if this means rejecting a generous offers and leaving both subjects with a $0 payoff. Similar rejections pattern had been previously reported in the field (Bahry and Wilson (2006) and Henrich et al. (2001)), but never in the lab or in a three-player setting. This result challenges some of the previous literature such as Fehr and Fischbacher (2004) or Falk et al. (2008), where third-parties reward generous offers and punish selfish ones.

To test the robustness of our results we introduce a number of variations to our original game. In a first variation we charge the decision-maker $1 if the game ends in a rejection; in a second one we substitute the proposer by a computer that randomly proposes a split of the $10. In both cases we continue to observe rejections of generous and selfish offers, and cannot find any statistical differences between the original treatment and the two variations. We, therefore, conclude that reducing absolute inequality[26] is the main concern of the decision-makers, while the intentions of the proposer play only a secondary role.

The above mentioned results could be worrisome for institutions relying on the decisions of neutral third-parties, since in our experiment not only do they make extremely inefficient decisions, but they also seem to ignore the intentions behind proposals. This latter finding, if general, could become a problem in our legal system where intentions and premeditation carry so much weight. And while it is beyond the scope of this paper to suggest a mechanism to correct the observed bias for equality in neutral third-parties, we believe that running further experiments in collaboration with faculty at Law Schools, or using subject pools composed by professional arbitrageurs or judges should be a natural next step. If such experiments confirmed our observations, in addition to inviting the appropriate institutional reforms, they would also no doubt promote the experimental method as a useful tool to improve legal regulation and institutions.

[26]As defined in section 4.1.

Appendix:

Appendix A: Details on session structure

The treatment ordering for each session as well as the total number of subjects per session in Table A.1

Treatment Order/Town	Barcelona	Santa Cruz
N2H	18	21
N2L	18	21
(H-1)2(L-1)	-	33
(L-1)2(H-1)	-	48
L2H	-	12
2NL	18	-
2NH	18	-
H2N	15	-
L2N	15	-
Lm2Hm	30	-
Hm2Lm	18	-

Table 11: Treatment ordering and number of B subject observations

In Table A.2 we present the total number of actual decision-maker observations for each treatment:

	Barcelona	Santa Cruz	Total
N	33	14	47
H	17	11	28
L	17	11	28
H-1	-	27	27
L-1	-	27	27
Lm	33	-	16
Hm	33	-	16

Table 12: Total number of B subject observations per treatment

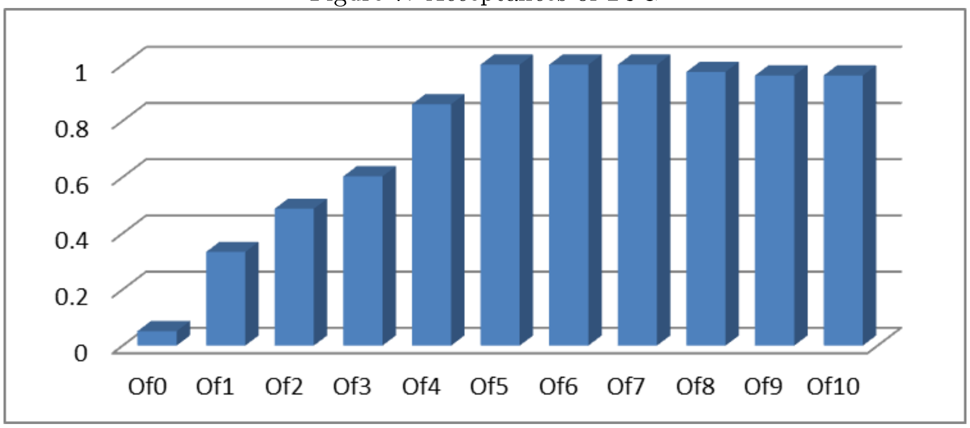
Appendix B: 2UG Results

We summarize all of B subject's observations in Figure 7. In it we present the percentage of decision-makers accepting each potential offer from A to C (e.g. almost 60% of B subjects accept a hypothetical offer of \$3 while only 30% accept one of 1). The acceptance results are slightly higher than those reported in the literature (see Camerer and Thaler (1995)), but still within the range of what would be expected. The average offer was of \$3.59, which is also what would be expected in an experiment like this. These results validate both our subject pool and the software interface, but most importantly, they show that decision-makers act consistently[27] when deciding

[27] Three subjects that rejected offers of \$8 or more yet accepted all smaller offers. We believe that these subjects misunderstood the interface and were trying to reject offers smaller than \$2.

about hyper-generous offers (i.e., subjects do not randomize or "experiment" within this range of offers). We take this as an indication that decision-makers take seriously the possibility of a generous offer.

Figure 7: Acceptances of 2UG



Appendix C: Two-sided Fisher test for baseline treatments

	\$0	\$1	\$2	\$3	\$4	\$5	\$6	\$7	\$8	\$9	\$10
L=H	1.000	0.775	0.596	1.000	1.000	0.141	0.550	1.000	1.000	0.810	1.000
H=N	0.355	0.280	0.202	0.808	0.604	0.250	0.759	0.792	0.226	0.469	0.636
L=H	0.329	0.227	0.089*	0.789	0.768	1.000	1.000	0.768	0.269	0.412	0.787

Table 13: Two-Sided Fisher P-values

Appendix D: Ordering Effects

Due to a miscommunication between the Barcelona and Santa Cruz labs we have a very unbalanced amount of for first round H treatment (5) compared with third round H treatment (22). This unfortunately pollutes the ordering effects for the H treatments as a 2 tailed Fisher Test comparing first round treatments against other rounds in the experiment shows.

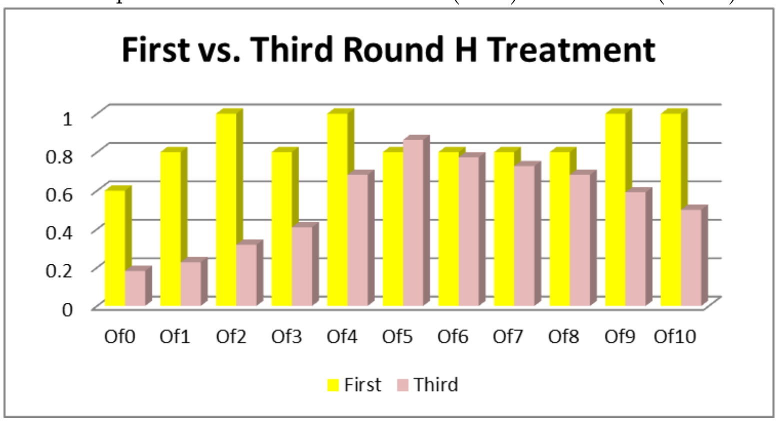
	\$0	\$1	\$2	\$3	\$4	\$5	\$6	\$7	\$8	\$9	\$10
N	0.752	0.890	0.344	0.671	0.174	1.000	0.767	0.492	0.357	0.923	0.628
H-1	0.704	1.000	1.000	0.090*	0.621	1.000	1.000	1.000	1.000	1.000	1.000
H	0.091*	0.030**	0.010**	0.165	0.238	1.000	1.000	1.000	1.000	0.136	0.060*
L	0.574	1.000	0.352	0.687	0.407	1.000	1.000	1.000	0.435	1.000	0.435
L-1	1.000	0.448	0.692	1.000	0.056*	0.549	0.549	1.000	0.662	0.662	0.448

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Two-Sided Fisher P-values Comparing First Round Treatments to all Other Treatments

While most treatments have no ordering effects, the LHT of the H treatment seems to be significantly affected by ordering. If we look at Graph A, we can see that while last round pattern of acceptances does look like those in the rest of treatments, first round H acceptances looks pretty random. As mentioned, we believe that this is due to the low number of observations of H in the first round, and that if we had more observations we would see no ordering effects.

Figure 8: Acceptance Rates for H for First (n=5) and Third (n=22) Round



Appendix E: Spearman Rank Correlation

In order to test for the correlation between distance and acceptance rates we first run a Spearman Rank Correlation test (Table 15) where a result of 1 or -1 is a perfect monotonic correlation of coefficients (in this case distance and acceptances). To run this test we divide our support into two separate tranches, the first one will include all offers to the left of $5 (LHT), the second tranche will include all offers to the right of five (RHT). As we can see, the RHT has a perfectly linear and highly significant relation between distance to the even split and acceptance levels; the closer to $5, the more acceptances we see. In the RHT the correlation is almost as perfect, in this case we see how as we get further away from $5 the levels of acceptance fall in a highly significant quasi-linear way.

Table 15: Spearman Rank Correlation Results for LHT and RHT under L, N and H treatments.

	LHT(L)	LHT(N)	LHT(H)	RHT(L)	RHT(N)	RHT(H)
Spearman Rho	1.000	1.000	1.000	-0.9411	-0.9429	-1.000
Prob > |t|	0.000	0.000	0.000	0.0051	0.0048	0.0000

Appendix F: Two-sided Fisher Test comparing same absolute inequality offers across all treatments in the baseline

Table 16: Two-sided Fisher Test.

Treatment	\$4=\$6	\$3=\$7	\$2=\$8	\$1=\$9	\$0=\$10
L	0.768	0.106	0.026**	0.011**	0.004***
H	1	0.093*	0.098*	0.029**	0.027**
N	0.048**	0.011**	0.006***	0.001***	0.0000***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix G: Spearman Rank Correlation

Table 17: Spearman Rank Correlation Results for LHT and RHT under L, N and H treatments.

	LHT(L-1)	LHT(H-1)	RHT(L-1)	RHT(H-1)
Spearman Rho	0.9856	1.000	-0.9710	-0.7495
Prob > |t|	0.0003	0.0000	0.0012	0.0059

Appendix H: Two-sided Fisher

	\$0	\$1	\$2	\$3	\$4	\$5	\$6	\$7	\$8	\$9	\$10
L-1 vs. H-1	1.000	0.782	0.779	1.000	1.000	0.610	1.000	0.467	1.000	1.000	0.224

Table 18: Two-Sided Fisher P-values Comparing First Round Treatments to all Other Treatments

Appendix I: One-sided Fisher Test for Machine Treatment

Table 19: One-sided Fisher P-values comparing total acceptances per treatment.

	\$0	\$1	\$2	\$3	\$4	\$5	\$6	\$7	\$8	\$9	\$10
L vs. Lm	0.434	0.456	0.026**	0.533	1.00	1.00	0.732	0.751	1.00	0.213	0.113
H vs. Hm	0.185	0.530	0.761	0.755	0.737	1.00	1.00	0.316	0.509	0.111	0.752

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix J: Instructions L2H

Welcome! This is an economics experiment. You will be a player in many periods of an interactive decision-making game. If you pay close attention to these instructions, you can earn a significant sum of money. It will be paid to you in cash at the end of the last period. It is important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and we will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your cooperation today.

This experiment has three different rounds. Before each round the specific rules and how you will earn money will be explained to you. In each round there will always be three types of players: A, B and C. You will be assigned to a type in Round 1 and will remain this type across all three rounds. Only one of the three rounds will be used for the final payoffs. This round is chosen randomly by the computer. The outcomes of each round are not made public until the end of the session (i.e. after round 3). Each round the groups are scrambled so you will never make offers or decide for the same player in two different rounds.

Round 1:

The first thing that you will see on your screen is your player type.

You will then be assigned to a group consisting of three players: an A type, B type and C type.

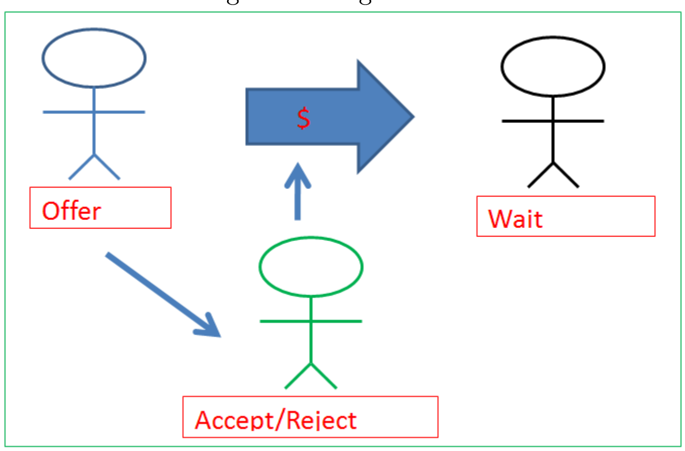
Player A will be endowed with $10 which he will split with player C. In order to do so Player A will have to input the amount he is willing to offer Player C. Player A will only be able to make integer offers (full dollars), so A will not be able to break its offer into cents.

While player A is deciding how much to offer player C, player B will be filling out a binding "strategy profile". The strategy profile has an "accept or reject" button for each potential offer from A to C (from $0 to $10). Player B's binding decision to accept or reject A's offers to C will be done before he knows the actual offer made by A.

A's decision: How to split an endowment of $10 with Player C by making him an offer between $0 and $10. If the offer is of $X, A will be keeping for himself 10-X.

B's decision: Before knowing the offer from A to C, B will fill a binding "strategy profile" deciding whether he accepts or rejects every potential offer from A to C. This decision is made without knowing the offer from A to C.

Figure 9: Diagram 3UG



It is very important for A to realize that he is going to write the amount he wants to offer C and not how much he wants to keep.

Payoff for Round 1:

If B accepts the offer from A to C, then they split the $10 as suggested by A.

If B rejects the offer from A to C, then both (A and C) get $0.

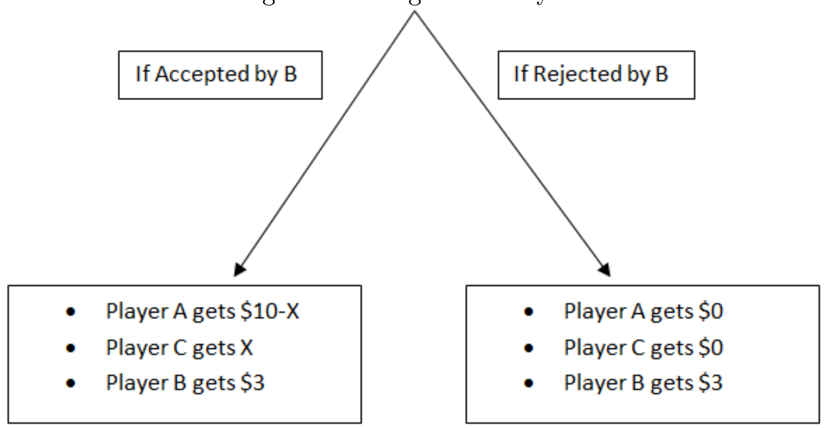
B will get paid $3 no matter what is the outcome.

Timing and Payoffs:

1. B fills a strategy profile with all potential offers from A to C.

2. A decides how much to offer C (say X)

Figure 10: Diagram of Payoffs



<u>Round 2</u>:

As mentioned at the beginning of the experiment you will keep your player type across the whole session. So A players are still A, B are B and C are C.

In this round type A players will be endowed with $20 and will have to make TWO offers:

1. How to split $10 with player B.

2. How to split $10 with player C.

As in Round 1 a binding "strategy profile" will be filled by B and C players before they know the offer made to *them.*

It is very important to notice that B and C players are making decisions concerning their own payoffs.

<u>A's decision</u>: How to split $10 with B and how to split $10 with A.

Each offer is independent. So the outcome of the offer to B has no effect on the outcome of the offer to C.

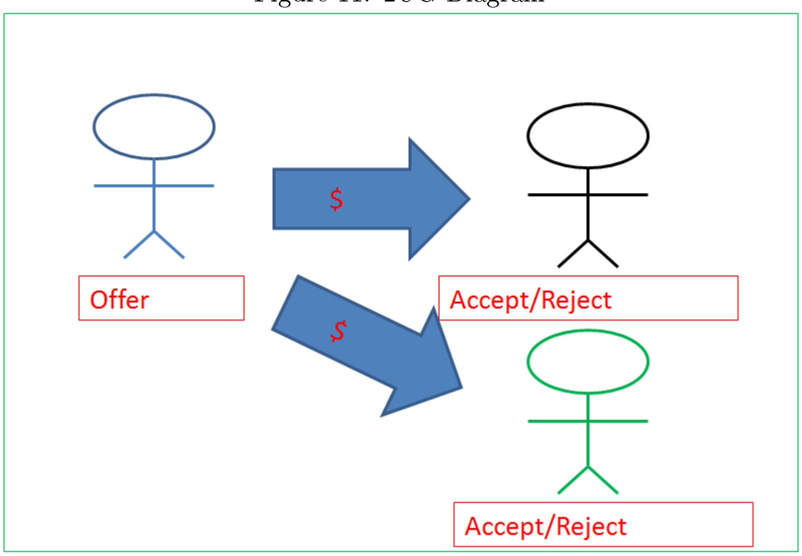
Payoffs for A will be as in Round 1 (if he offers X and the offer is accepted he gets $10-X, if the offer is rejected both him and the rejecting player get 0).

B and C players will get paid X or 0 depending if the accepted or rejected the offer made directly to them.

In order to make payoffs equitable for this round, A's payoff for this round will be chosen at random between one of the two outcomes (offer to B and offer to C). B and C's decision: Before knowing the offer made to them by Player A, B and C will fill a binding "strategy profile" deciding if they accept or reject *every potential offer made directly to them.*

If the offer from A is accepted, then the split is done as proposed by A. If the offer is rejected both the receiver and A get $0 as the outcome for this round.

Figure 11: 2UG Diagram



<u>Timing and Payoff for Round 2</u>:

1. Each receiver fills a strategy profile with all potential offers that A could make them.

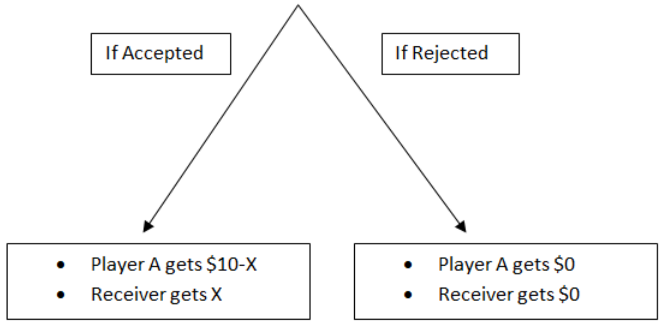
2. A decides how much to offer C and B (say X)

3. Payoffs for B and C will be the outcome of their particular game with A.

4. To make outcomes equitable, the computer will choose randomly one of the two outcomes to be A's payoff for the round.

For each offer made from A to the other members of his group:

Figure 12: 2UG Payoffs



Round 3:

As mentioned at the beginning of the experiment you will remain your player type across the whole session.

This round is very similar to round 1. You will now be re-scrambled into groups of three subjects (one A, one B and one C subject).

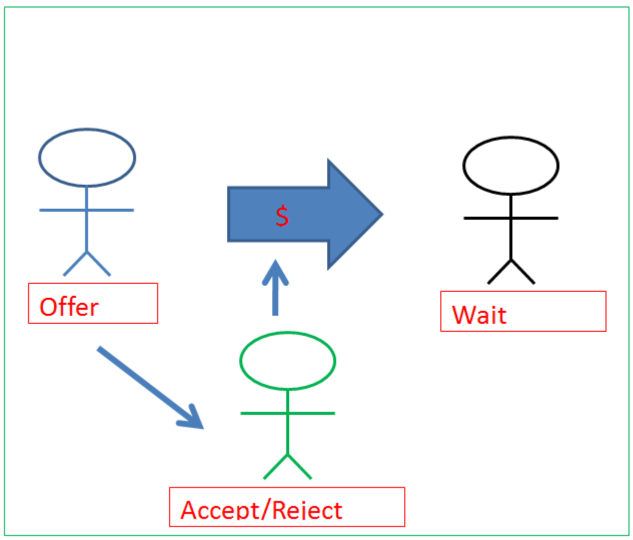
A will be endowed with $10 and must decide how to split them with C.

B's role is exactly the same as that in round 1: Before knowing the offer from A to C, B will fill a "strategy profile" deciding whether he accepts or rejects *every potential offer from A to C*.

If the offer from A to C is accepted by B, then the split is done as proposed by A. If B rejects the offer, then both A and C receive $0 for this round.

B's payoff in this round is a flat $12 fee, whatever his decision and outcome of the round. So, the only change between Round 1 and Round 3 is that player B, is getting paid a different amount.

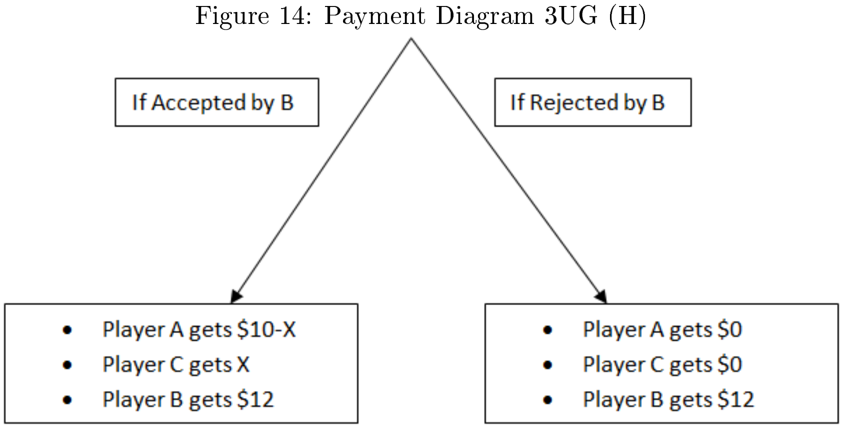
Figure 13: 3UG (H) Diagram



Timing and Payoffs:

1. B fills a strategy profile with all potential offers from A to B.
2. A decides how much to offer C (say X)

Figure 14: Payment Diagram 3UG (H)



References

Bahry, Donna L. and Rick K. Wilson, "Confusion or fairness in the field? Rejections in the ultimatum game under the strategy method," *Journal of Economic Behavior & Organization*, May 2006, *60* (1), 37–54.

Blount, Sally, "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences," *Organizational Behavior and Human Decision Processes*, 1995, *63* (2), 131–144.

Brandts, Jordi and Gary Charness, "Hot vs. cold: Sequential responses and preference stability in experimental games," *Experimental Economics*, 2000, *2* (3), 227—238.

_ and _ , "The strategy versus the direct-response method: a first survey of experimental comparisons," *Experimental Economics*, September 2011, *14* (3), 375–398.

Camerer, Colin and Richard H Thaler, "Anomalies: Ultimatums, Dictators and Manners," *Journal of Economic Perspectives*, May 1995, *9* (2), 209–219.

Charness, Gary and Matthew Rabin, "Understanding social preferences with simple tests," *The Quarterly Journal of Economics*, 2002, *117* (3), 817—869.

Cremin, Hilary, *Peer mediation: citizenship and social inclusion in action*, Maidenhead: Open University Press, 2007.

Ehrenberg, Victor, *From Solon to Socrates: Greek history and civilization during the sixth and fifth centuries B.C.*, London [etc.]: Routledge, 1973.

Engelmann, Dirk and Martin Strobel, "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments," *The American Economic Review*, 2004, *94* (4), 857—869.

Falk, Armin, Ernst Fehr, and Urs Fischbacher, "Driving forces behind informal sanctions," *Econometrica*, 2005, *73* (6), 2017—2030.

_ , _ , and _ , "Testing theories of fairnes–Intentions matter," *Games and Economic Behavior*, 2008, *62* (1), 287—303.

Fehr, Ernst and Klaus M Schmidt, "A theory of fairness, competition, and cooperation," *The quarterly journal of economics*, 1999, *114* (3), 817—868.

_ and Urs Fischbacher, "Third-party punishment and social norms," *Evolution and Human Behavior*, March 2004, *25* (2), 63–87.

Greiner, Ben, "An Online Recruitment System for Economic Experiments," 2004.

Güth, Werner and Eric van Damme, "Information, Strategic Behavior, and Fairness in Ultimatum Bargaining: An Experimental Study," *Journal of Mathematical Psychology*, June 1998, *42* (2–3), 227–247.

_ , Carsten Schmidt, and Matthias Sutter, "Bargaining outside the lab-a newspaper experiment of a three-person ultimatum game*," *The Economic Journal*, 2007, *117* (518), 449—469.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath, "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies," *The American Economic Review*, May 2001, *91* (2), 73–78. ArticleType: research-article / Issue Title: Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association / Full publication date: May, 2001 / Copyright © 2001 American Economic Association.

Kagel, John H and Katherine Willey Wolfe, "Tests of fairness models based on equity considerations in a three-person ultimatum game," *Experimental Economics*, 2001, *4* (3), 203—219.

Knez, Marc J. and Colin F. Camerer, "Outside Options and Social Comparison in Three-Player Ultimatum Game Experiments," *Games and Economic Behavior*, July 1995, *10* (1), 65–94.

Leibbrandt, Andreas and Raul Lopez-Perez, "The envious punisher," Technical Report, University of Zurich working paper 2008.

Martin, Jonathan Law Elizabeth A., *A Dictionary of Law*, Oxford University Press, 2003.

Ockenfels, Axel and Gary E. Bolton, "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, March 2000, *90* (1), 166–193.

Turnuklu, Abbas, Tarkan Kacmaz, Dilara Sunbul, and Hatice Ergul, "Does peer-mediation really work? effects of conflict resolution and peer-mediation training on high school students' conflicts," *Procedia - Social and Behavioral Sciences*, 2009, *1* (1), 630–638.

Xiao, Erte and Daniel Houser, "Emotion expression in human punishment behavior," *Proceedings of the National Academy of Sciences of the United States of America*, 2005, *102* (20), 7398—7401.

SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Principal Component Analysis in an Asymmetric Norm" by Ngoc Mai Tran, Maria Osipenko and Wolfgang Karl Härdle, January 2014.
002 "A Simultaneous Confidence Corridor for Varying Coefficient Regression with Sparse Functional Data" by Lijie Gu, Li Wang, Wolfgang Karl Härdle and Lijian Yang, January 2014.
003 "An Extended Single Index Model with Missing Response at Random" by Qihua Wang, Tao Zhang, Wolfgang Karl Härdle, January 2014.
004 "Structural Vector Autoregressive Analysis in a Data Rich Environment: A Survey" by Helmut Lütkepohl, January 2014.
005 "Functional stable limit theorems for efficient spectral covolatility estimators" by Randolf Altmeyer and Markus Bibinger, January 2014.
006 "A consistent two-factor model for pricing temperature derivatives" by Andreas Groll, Brenda López-Cabrera and Thilo Meyer-Brandis, January 2014.
007 "Confidence Bands for Impulse Responses: Bonferroni versus Wald" by Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, January 2014.
008 "Simultaneous Confidence Corridors and Variable Selection for Generalized Additive Models" by Shuzhuan Zheng, Rong Liu, Lijian Yang and Wolfgang Karl Härdle, January 2014.
009 "Structural Vector Autoregressions: Checking Identifying Long-run Restrictions via Heteroskedasticity" by Helmut Lütkepohl and Anton Velinov, January 2014.
010 "Efficient Iterative Maximum Likelihood Estimation of High-Parameterized Time Series Models" by Nikolaus Hautsch, Ostap Okhrin and Alexander Ristig, January 2014.
011 "Fiscal Devaluation in a Monetary Union" by Philipp Engler, Giovanni Ganelli, Juha Tervala and Simon Voigts, January 2014.
012 "Nonparametric Estimates for Conditional Quantiles of Time Series" by Jürgen Franke, Peter Mwita and Weining Wang, January 2014.
013 "Product Market Deregulation and Employment Outcomes: Evidence from the German Retail Sector" by Charlotte Senftleben-König, January 2014.
014 "Estimation procedures for exchangeable Marshall copulas with hydrological application" by Fabrizio Durante and Ostap Okhrin, January 2014.
015 "Ladislaus von Bortkiewicz - statistician, economist, and a European intellectual" by Wolfgang Karl Härdle and Annette B. Vogt, February 2014.
016 "An Application of Principal Component Analysis on Multivariate Time-Stationary Spatio-Temporal Data" by Stephan Stahlschmidt, Wolfgang Karl Härdle and Helmut Thome, February 2014.
017 "The composition of government spending and the multiplier at the Zero Lower Bound" by Julien Albertini, Arthur Poirier and Jordan Roulleau-Pasdeloup, February 2014.
018 "Interacting Product and Labor Market Regulation and the Impact of Immigration on Native Wages" by Susanne Prantl and Alexandra Spitz-Oener, February 2014.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

019 "Unemployment benefits extensions at the zero lower bound on nominal interest rate" by Julien Albertini and Arthur Poirier, February 2014.
020 "Modelling spatio-temporal variability of temperature" by Xiaofeng Cao, Ostap Okhrin, Martin Odening and Matthias Ritter, February 2014.
021 "Do Maternal Health Problems Influence Child's Worrying Status? Evidence from British Cohort Study" by Xianhua Dai, Wolfgang Karl Härdle and Keming Yu, February 2014.
022 "Nonparametric Test for a Constant Beta over a Fixed Time Interval" by Markus Reiß, Viktor Todorov and George Tauchen, February 2014.
023 "Inflation Expectations Spillovers between the United States and Euro Area" by Aleksei Netšunajev and Lars Winkelmann, March 2014.
024 "Peer Effects and Students' Self-Control" by Berno Buechel, Lydia Mechtenberg and Julia Petersen, April 2014.
025 "Is there a demand for multi-year crop insurance?" by Maria Osipenko, Zhiwei Shen and Martin Odening, April 2014.
026 "Credit Risk Calibration based on CDS Spreads" by Shih-Kang Chao, Wolfgang Karl Härdle and Hien Pham-Thu, May 2014.
027 "Stale Forward Guidance" by Gunda-Alexandra Detmers and Dieter Nautz, May 2014.
028 "Confidence Corridors for Multivariate Generalized Quantile Regression" by Shih-Kang Chao, Katharina Proksch, Holger Dette and Wolfgang Härdle, May 2014.
029 "Information Risk, Market Stress and Institutional Herding in Financial Markets: New Evidence Through the Lens of a Simulated Model" by Christopher Boortz, Stephanie Kremer, Simon Jurkatis and Dieter Nautz, May 2014.
030 "Forecasting Generalized Quantiles of Electricity Demand: A Functional Data Approach" by Brenda López Cabrera and Franziska Schulz, May 2014.
031 "Structural Vector Autoregressions with Smooth Transition in Variances – The Interaction Between U.S. Monetary Policy and the Stock Market" by Helmut Lütkepohl and Aleksei Netsunajev, June 2014.
032 "TEDAS - Tail Event Driven ASset Allocation" by Wolfgang Karl Härdle, Sergey Nasekin, David Lee Kuo Chuen and Phoon Kok Fai, June 2014.
033 "Discount Factor Shocks and Labor Market Dynamics" by Julien Albertini and Arthur Poirier, June 2014.
034 "Risky Linear Approximations" by Alexander Meyer-Gohde, July 2014
035 "Adaptive Order Flow Forecasting with Multiplicative Error Models" by Wolfgang Karl Härdle, Andrija Mihoci and Christopher Hian-Ann Ting, July 2014
036 "Portfolio Decisions and Brain Reactions via the CEAD method" by Piotr Majer, Peter N.C. Mohr, Hauke R. Heekeren and Wolfgang K. Härdle, July 2014
037 "Common price and volatility jumps in noisy high-frequency data" by Markus Bibinger and Lars Winkelmann, July 2014
038 "Spatial Wage Inequality and Technological Change" by Charlotte Senftleben-König and Hanna Wielandt, August 2014
039 "The integration of credit default swap markets in the pre and post-subprime crisis in common stochastic trends" by Cathy Yi-Hsuan Chen, Wolfgang Karl Härdle, Hien Pham-Thu, August 2014

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2014

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

040 "Localising Forward Intensities for Multiperiod Corporate Default" by Dedy Dwi Prastyo and Wolfgang Karl Härdle, August 2014.
041 "Certification and Market Transparency" by Konrad Stahl and Roland Strausz, September 2014.
042 "Beyond dimension two: A test for higher-order tail risk" by Carsten Bormann, Melanie Schienle and Julia Schaumburg, September 2014.
043 "Semiparametric Estimation with Generated Covariates" by Enno Mammen, Christoph Rothe and Melanie Schienle, September 2014.
044 "On the Timing of Climate Agreements" by Robert C. Schmidt and Roland Strausz, September 2014.
045 "Optimal Sales Contracts with Withdrawal Rights" by Daniel Krähmer and Roland Strausz, September 2014.
046 "Ex post information rents in sequential screening" by Daniel Krähmer and Roland Strausz, September 2014.
047 "Similarities and Differences between U.S. and German Regulation of the Use of Derivatives and Leverage by Mutual Funds – What Can Regulators Learn from Each Other?" by Dominika Paula Gałkiewicz, September 2014.
048 "That's how we roll: an experiment on rollover risk" by Ciril Bosch-Rosa, September 2014.
049 "Comparing Solution Methods for DSGE Models with Labor Market Search" by Hong Lan, September 2014.
050 "Volatility Modelling of CO2 Emission Allowance Spot Prices with Regime-Switching GARCH Models" by Thijs Benschopa, Brenda López Cabrera, September 2014.
051 "Corporate Cash Hoarding in a Model with Liquidity Constraints" by Falk Mazelis, September 2014.
052 "Designing an Index for Assessing Wind Energy Potential" by Matthias Ritter, Zhiwei Shen, Brenda López Cabrera, Martin Odening, Lars Deckert, September 2014.
053 "Improved Volatility Estimation Based On Limit Order Books" by Markus Bibinger, Moritz Jirak, Markus Reiss, September 2014.
054 " Strategic Complementarities and Nominal Rigidities" by Philipp König, Alexander Meyer-Gohde, October 2014.
055 "Estimating the Spot Covariation of Asset Prices – Statistical Theory and Empirical Evidence" by Markus Bibinger, Markus Reiss, Nikolaus Hautsch, Peter Malec, October 2014.
056 "Monetary Policy Effects on Financial Intermediation via the Regulated and the Shadow Banking Systems" by Falk Mazelis, October 2014.
057 "A Tale of Two Tails: Preferences of neutral third-parties in three-player ultimatum games" by Ciril Bosch-Rosa, October 2014.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".