A Tale of Two Tails:

Rejection Patterns of Extreme Offers in a Three Player Game

Ciril Bosch-Rosa
University of California at Santa Cruz
(cboschro@ucsc.edu)

January 17, 2011

Abstract: We present a three-player game, in which a decision-maker, in the role of referee, accepts or rejects the offer made by a proposer to a passive receiver. The results show a high level of rejection of both selfish and generous offers by the referee. We show that contrary to the best-known models of social preferences, our judge's decisions are independent of their payoff. In addition, we are able to show that concerns for intentions of proposers are secondary when compared to inequality concerns.

JEL: C92, D71, D63, D31
Keywords: inequality, ultimatum game, fairness, intentions, experiment, hyper-fair

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. Thanks also to Johannes Müller-Trede for running some of the Barcelona experiments and for endless discussion, and 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 seminar were of great help.

> "And the king said, Divide the living child in two,
> and give half to the one, and half to the other."
> Kings 3:25

1. Introduction

In this paper we present the results of a three-player game involving a proposer, who suggests how to split $10 between him and a passive receiver, while a referee fills a strategy profile accepting or rejecting every potential offer by the proposer. If the offer is accepted, then the split takes place as suggested by the proposer. If it is rejected, then both the proposer and the passive responder get $0. The referee's payoff depends on the treatment.

These treatments will be divided into two different families. In the first one, called "free-rejection", referees get paid a flat-fee regardless of the outcome of the game. In the second one, referees have to pay $1 of their flat-fee if the game ends in a rejection. Within both families we will have high and low treatments.

Our results show a referee that is extremely concerned with equality, willing to reject both generous and selfish offers. We will see how acceptance always peaks at the fair split ($5) and then monotonically drops *in both directions* as offers become more unequal. Indeed, like King Salomon in the famous trial, our referees prefer to "divide the child" rather than end up with an unfair outcome.

We will see how in our setup referees *do not* take their payoff as a reference point from which to judge verdicts, something that runs against the most known inequality aversion models such as Fehr and Schmidt (1999) or Bolton and Ockenfelds (2000). Another finding will be that concerns for equality are much stronger than those for intentions in the costly-rejection family.

The above mentioned outcomes of our analysis could be a combination of good and bad news for our judicial system. While it is good news that people make their resolutions independent of their payoffs, the tendency that they show for Pareto-inefficient outcomes is worrisome. After all, Salomon did have an ulterior goal that we do not believe our subjects have.

2. Literature review

The setup we described is very similar to some of the literature on three-player ultimatum games, a line of research that has been studied as far back as Knez and Camerer (1995). In this experiment, a proposer would make two offers, one for each responder, while these made an acceptance decision conditional on the offer made to the other subject. The results showed that subjects used the other player's offer as a reference point to accept or reject an identical proposal. In Güth and Van Damme (1998) a proposer had to split the whole pie with a decision maker and a dummy player; if the offer was accepted, then the split went as suggested, if rejected, then everyone received zero. The outcome was that both proposer and responder ended up giving a minimal share to the dummy player, and splitting most of it amongst themselves. Finally, Kagel & Wolfe (2001) present us with a setup identical to Güth and Van Damme

(1998) though now, if the offer was rejected, the dummy player would get a consolation prize. As in Güth and Van Damme (1998), the authors observe how the dummy player is completely ignored.

In literature on third–party costly punishment, Fehr, Falk and Fischbacher (2005) report that sanctions are not used to reduce inequality, but rather as retaliation for selfish actions. On the other hand Leibbrandt and Lopez-Perez (2008) use a within-subject structure to conclude that both 2nd and 3rd party punishments are driven by outcome not intention. Interestingly, and against Fehr and Fischbacher (2004), Leibbrandt and Lopez-Perez (2008) find that 2nd party punishment is not significantly higher than 3rd party punishment. More recently Falk *et al*. (2008) have revisited the subject suggesting that while inequality has some effect on punishment, intentions are the main reason behind punitive actions.

As we mentioned, one of our results is a pattern of acceptance with rejection of both selfish and generous offers. To our knowledge this result had been observed previously in the field, but never in the lab. Further, when observed it was always dismissed as an anomaly. Bahry and Wilson (2005) report an "inverted-U" pattern of acceptances when comparing the ultimatum game results they achieved in old Soviet Union regions. Their results are similar to ours in rejections of what they call "hyper-fair" offers, but concluded that this result was due to a Soviet education and dismissed it. To the extent of our knowledge, subjects involved in our experiments have not been raised with socialist ideals, so any rejection of hyper-fair offers cannot be blamed on a Marxist education. The second case of a paper reporting Inverted-U patterns is Güth *et al*. (2007), which gathered ultimatum game data through newspaper publications. Even if the authors recorded this type of behavior, they dismissed it due to the insignificant number of cases.

Finally, to validate the use of the strategy method, we refer to Brandts and Charness (2010). In this survey the authors come to the conclusion that there does not seem to be an overall effect of "hot vs. cold" decisions. In fact, in those cases where there seem to be differences, using a direct method would most probably accentuate or inverted-U results. Further, as reported in the paper, "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 (2003) for more information on the subject.

3. Experimental Design

The experiment was run with a total of 233 undergraduates from both Universitat Pompeu Fabra (UPF) in Barcelona and the University of California at Santa Cruz (UCSC). Each session had three rounds with different games and treatments, and overall they lasted around 25 minutes, with average earnings of $3.5 at UCSC and €3.35 at UPF[1]. Subjects were recruited through the internal electronic systems of each university with the only requirement being no previous experience in bargaining games. In total 15 sessions were run, UCSC sessions had 12 subjects[2] and those at UPF 18[3].

[1] On top of this, subjects had a show-up fee of $5 in Santa Cruz or 3€ in Barcelona which was announced only at the end of the experiment.
[2] Except 3 sessions that had 9 subjects.
[3] Except 2 sessions that had 12 subjects.

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 would have three rounds and instructions for each of these would be read immediately before the round started.
2) Each subject would be assigned a player type (A, B or C) which they would keep throughout the experiment.
3) Each round subjects would be randomly assigned to a group of three players (one of each type) to avoid repeated game effects.
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 game, when they would be informed of the actions of every (anonymous) subject in their group and the random payoff the computer had chosen.

Each round subjects would play one of the two games we designed; a three-player ultimatum game (3UG) or a regular ultimatum game (2UG) which we used as control.

3.1 3UG:

The three-player ultimatum game has a proposer (A subject) making an offer to a dummy player (C subject) on how to split $10 (€10 in Barcelona). Meanwhile decision-makers (B subjects) fill a strategy profile (Figure 1) accepting or rejecting all potential offers from A to C. If the offer is accepted, then the split goes as suggested by A, if rejected then both A and C get nothing for the round. B's payoffs on the other hand depend only on the treatment for each round.

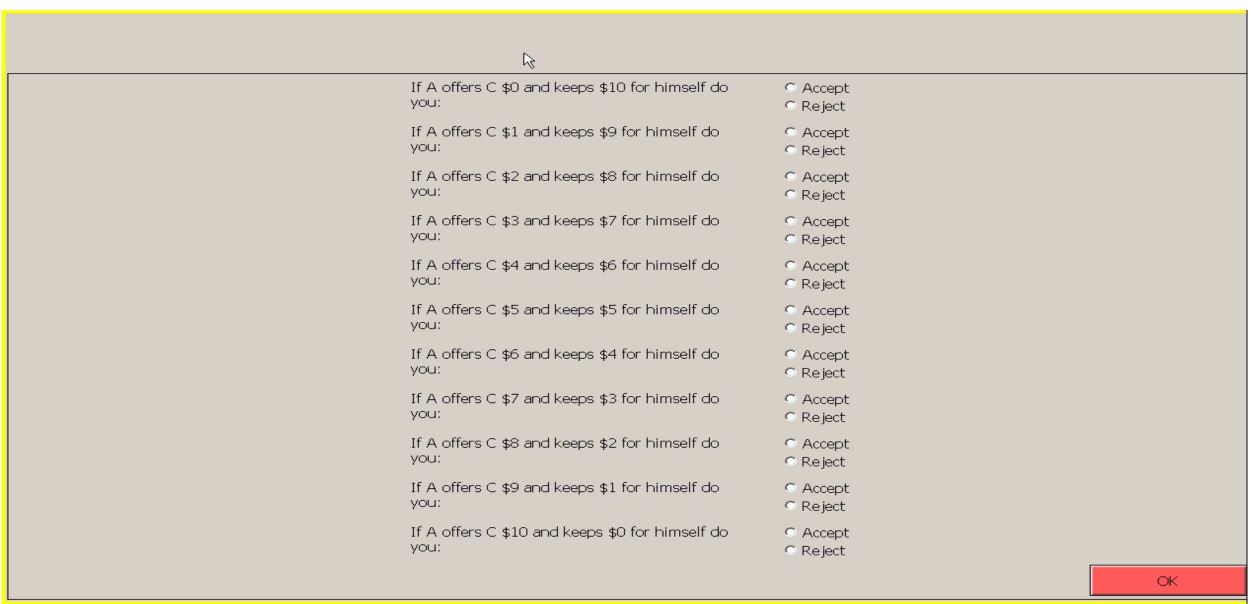

Figure 1: Decision-Maker Screenshot

Under the free-rejection family B subjects get paid a flat-fee independent of the outcome of the game. There are three different treatments in this family:

- Normal (N): B gets paid $5 (€5) for his decisions.
- Low (L): B gets paid $3 (€3) for his decisions.
- High (H): B gets paid $12 (€12) for his decisions.

The costly-rejection family follows the same structure, except that there is a penalty of $1 for B players if the game ends in a rejection. The treatments are[4]:

- 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.

The layout of the game in both families can be seen in Figure 2:

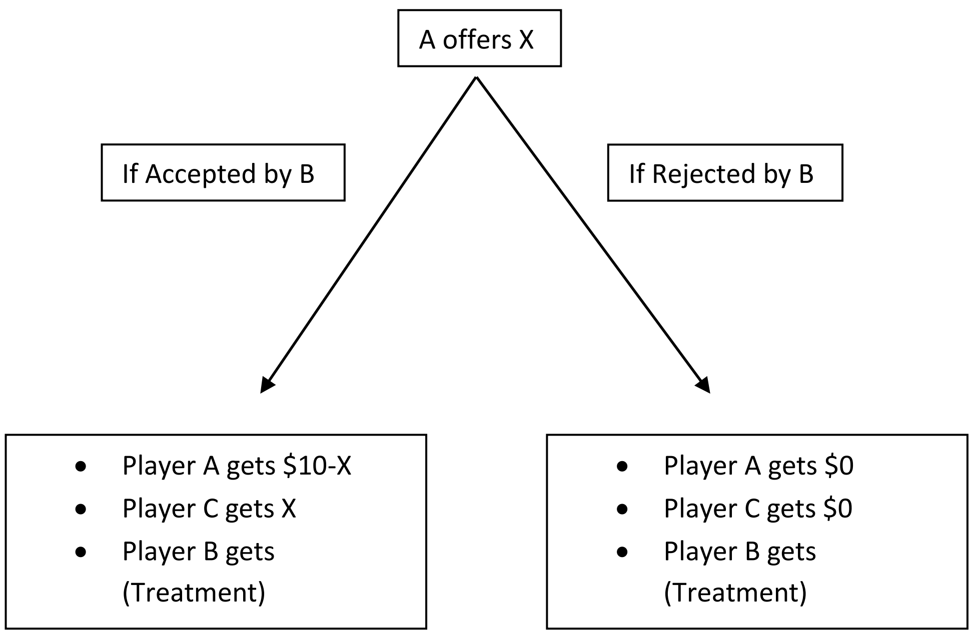

Figure 2: General Structure of the Experiment

3.2 2UG:

In this game, A makes two independent offers; one to B and the other to C. These offers are on how to split two independent pies of $10 (€5). As in the 3UG case, both B and C fill a strategy profile, accepting or rejecting all potential offers from A *to them*.

If B (C) rejects the offer, then B (C) gets $0 for the round; if accepted, then the split goes as suggested. A on the other hand gets paid only one of the two outcomes (either that of the offer made to B or that of

[4] These types of sessions were only run at UCSC, so payments are only in dollars.

the offer made to C). If it turns out to be a rejection, then A gets $0, if an acceptance, then A gets his part of the proposal. This randomization of payoffs is done in order to prevent portfolio effects, and to make payoffs fairer across subject types.

The 2UG game was introduced to control; if the results were similar to those of regular ultimatum games, then this would validate both our population and game interface.

3.3 Structure of the sessions:

The ordering of the sessions and the amount of total subjects in each session is found in Table 1:

Treatment Order	**Number of A, B and C subjects per Treatment**	
	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	-

Table 1: Order of Treatments and Number of Subjects Participating

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

	B Subject Observations by 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

Table 2: Total Observations per Treatment and Localization

4. Results:

4.1 2UG game:

In total we have 78 observations (we ignore C subject's answers) which we summarize in Figure 3, which shows the percentage of acceptances for each offer (e.g. almost 60% of subjects accept an offer of $3[5] and 30% an offer of $1). The acceptance results are perhaps slightly higher than those reported on average in the literature (Camerer and Thaler (1995)), but still whithin the range of what we would expect. The average offer was of $3.59, which is also what would be expected of an experiment like this.

[5] From now on we will use the dollar sign as a generalization for euros and dollars.

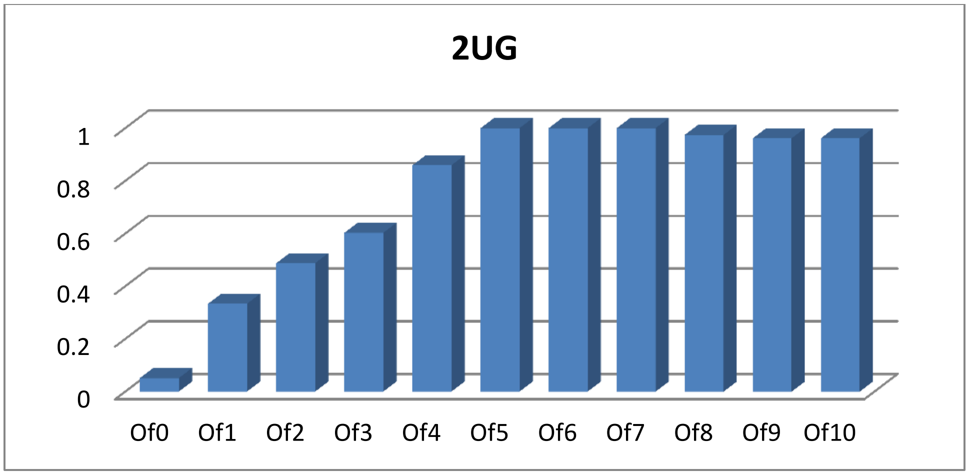

Figure 3: Acceptance rates 2UG treatment

Given that our 2UG results are typical of the published literature, we can safely attribute the results of 3UG to treatments to the structure of our game rather than to our pool of subjects. Furthermore, and important for our future analysis, Figure 3 shows that subjects do believe that generous offers are possible, and typically don't respond to generous offers in an inconsistent or random way[6].

4.2 3UG game:

4.2.1 Free-Rejection Family:

In order to make the analysis easier to understand, we present in Figure 4 the results and comparisons of all treatments in the family. In the upper left corner we see the our baseline treatment N and in a clockwise order the comparison between N and L, L and H, and finally that between N and H in the lower left quadrant.

Two things stand out immediately from these graphs. The first one is that under all treatments we see an inverted-U pattern peaking at the fair split ($5), the second striking feature is the similarity between treatments. Both these results were unexpected.

In order to study the outcome of our sessions we will adopt a structure of posing a question and then answering it.

[6] There are three subjects that indeed reject offers of $8 or more yet accept all smaller offers. We believe that these subjects misunderstood the instructions and were in fact trying to reject offers smaller than $2.

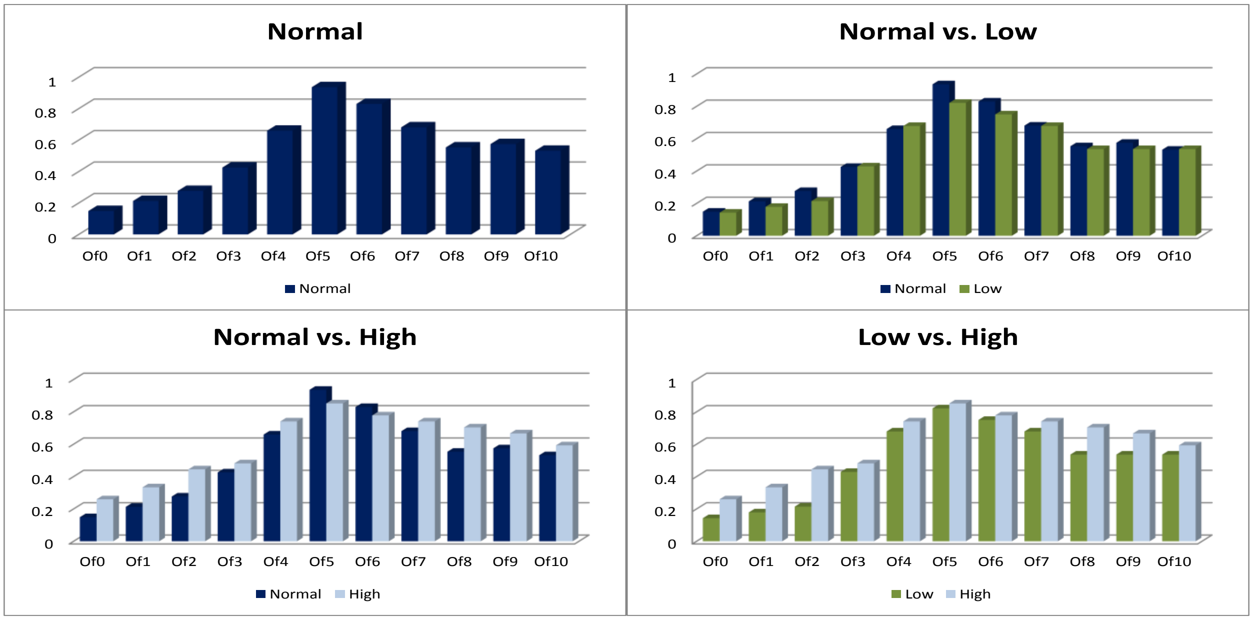

Figure 4: Acceptance Rate Histograms for all Three Free Rejection Treatments, with Pairwise Comparisons

Question 1: *Do differences in flat- fee payoffs have any effect on decision-makers?*

To test if flat-fee payments have an effect on total number of accepted offers we run a Wilcoxon Matched-pairs Signed-rank Test in which N is our baseline. The results shows no difference in total acceptances per subject between the N and L treatments (p-value=0.375) nor across N and H treatments (p-value=0.161) [7].

To reinforce our results we run a regression of total accepted offers (total) on dummies for location (where)[8], order (first) and treatment (high and low). The results are in Table 3; in the first two columns we compare High to Low, while in the third and fourth column we compare High and Low to the baseline N.

As we can see, neither when we compare H to L or both of these to N do we see any significant effect of the treatment dummies. Neither can we see any ordering or location effects when comparing H and L to N, something that is reconfirmed later in Table 6. Special mention should be made of Column 2 where we see some ordering effects that are due to the lack of first round H treatment observations (See Appendix A for a lengthier discussion).

[7] On the other hand the test becomes more significant when comparing L and H (p-value=0.0825), but recall that as mentioned above the number of observations comparing H and L is extremely low n=4.

[8] where=1 if session was run in Santa Cruz.

	(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)
N	55	55	102	102
adj. R^2	0.006	0.014	-0.004	-0.008

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Regression of total accepted offers by subject and treatment

Given that neither order nor location has any effects, we decide to run a Two-sided Fisher Test[9] comparing the aggregated number of acceptances for each individual. Table 4 shows how acceptance rates are not significantly different across treatments except for offers of $2 between H and L.

P-values	$0	$1	$2	$3	$4	$5	$6	$7	$8	$9	$10
L = N	1	.775	.596	1	1	.141	.55	1	1	.81	1
H = N	.355	.280	.202	.808	.604	.250	.759	.792	.226	.469	.636
L = H	.329	.227	.089*	.789	.768	1	1	.768	.269	.412	.787

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

Table 4: Two-sided Fisher P-Values Comparing Total Acceptances Across Treatments (* p < 0.10, ** p < 0.05, *** p < 0.01).

With table 3 and 4 we believe that we can answer Question 1 in a negative way. In the free-rejection family flat-fees do not make a difference, as B subjects make decisions independently of the payoff.

Answer 1: *No, flat fee payoffs have no significant effect on decision-maker's acceptance behavior.*

Given that decision-makers don't seem to take into account their own payoffs, we want to see if inequality between other participants is a reason to reject offers. To test this we look at relationship between distance to $5 of an offer and probability of acceptance.

Question 2: *Does the probability of acceptance have a negative correlation with distance to the even split ($5)?*

[9] A One-sided Test would not change the results much.

First we run a Spearman Rank Correlation Test between distance and acceptance rates to test this idea. To do so we divide our support into two separate groups; all of the offers to the right of $5, the Right Hand Tail (RHT), and all offers to the left of $5, the Left Hand Tail (LHT). As we can see in Table 5, the RHT has a perfectly linear and highly significant relation between distance to the even split and acceptance levels (i.e. the closer an offer is to $5, the more acceptances it gets). In the RHT we see the same; as we get further away from $5 the levels of acceptance fall in a highly significant and practically linear way.

	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	0.000

Table 5: Spearman Rank Correlation Results for LHT and RHT under L, N and H treatments.

In order to have a more accurate idea of how distance to the even split affects the probability of rejection, we run a linear probability model (Table 6). In it we have the binary accept/reject outcome as our dependent variable (accept in the table) and a series of dummies: ordering (first), treatment (high, low) and location (where), along with dummies for distance, which are coded with the distance to $5 and the tail in which they are located. For example, dist3l is the dummy for the $2 offer (3 dollars to the left of $5) and dist2r is the dummy for an offer of $7 (2 dollars to the right of $5).

As we can see in Column 5, all dummies for distance are highly significant and negative. Furthermore, if we go down the column we can see how the further an offer is from $5 the lower the probability of acceptance. These probabilities follow a monotonic drop in both tails[10], ranging from a 20% lower probability of acceptance for an offer of $4 (dist1l) to a 70% lower probability of acceptance for an offer of $0 (dist5l).

The other interesting result from Table 6 is the difference in coefficients of equidistant offers. Even if they are all negative and highly significant, there is a big difference in acceptance probabilities if the offer is generous (RHT) than if it selfish (LHT). For example; even if the level of inequality amongst A and C subjects is the same with an offer of $3 than with one of $7, their acceptance probabilities compared to a 50/50 split are vastly different (-44% and -19% respectively).

[10] Strictly monotonic in the LHT and weakly in the RHT.

	(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		0.0917	0.0917	0.0917	0.0917
		(0.0581)	(0.0582)	(0.0582)	(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)		(0.0584)
dist1r				0.340^{***}	-0.088^{***}
				(0.0447)	(0.0306)
dist2r				0.242^{***}	-0.186^{***}
				(0.0485)	(0.0467)
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)
N	1122	1122	1122	1122	1122
adj. R^2	0.004	0.008	0.164	0.055	0.193

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Linear Probability Model of Acceptances

It seems apparent from our results that inequality amongst subjects plays a key role in the decision-makers accept/reject decisions.

Answer 2: *Yes, the further away an offer is from the even split ($5), the lower the probability of it being accepted.*

Our results in Table 6 show that the tails are not symmetric. Our interpretation is that this is due to decision-makers punishing selfish (rewarding generous) proposers and thus breaking the balance of the distribution. This takes us to Question 3.

Question 3: *Are LHT and RHT symmetric?*

When answering this question we will look at the tails, but only *within treatment*. To do this we run a Wald Test comparing the coefficients of a probabilistic model such as that in column 5 of Table 6. The results show that equidistant dummies are significantly different in all treatments (Table 7).

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***

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

Table 7: P-values of Wald Test, Testing for Equality in Within Treatment Regression Coefficients

For completeness in Table 8 we show the results of a Two-Sided Fisher Test comparing the number of accepted offers, with practically identical results.

Treatment	$4 = $6	$3 = $7	$2 = $8	$1 = $9	$0 = $10
L	0.768	0.106	0.026**	0.011**	0.004***
H	1.00	0.093*	0.098*	0.029**	0.027**
N	0.048**	0.011**	0.006***	0.001***	0.000***

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

Table 8: Two-sided Fisher P-values Comparing Equidistant Offers.

Both these results confirm our suspicion that decision-makers are concerned with intentions and willing to reject selfish offers, even if it is at the cost of the (small) payoff of the dummy player.

Answer 3: *No, within treatment, equidistant offers have different acceptance probabilities depending on the tail they are.*

We have analyzed the results of our three-player ultimatum game by answering three questions about how decision-makers filled their strategy profile. Our first conclusion is that our “referees” do not take into account their own payoff when making decisions concerning other subjects payoffs. Unfortunately, this comes at the cost of many Pareto-inefficient decisions being made (including rejection of generous offers!). On the other hand, we have been able to show that even some of these rejections are motivated by the intentions of the proposer.

4.2.2 Costly Rejection:

In view of our results we introduce the costly-family which has treatments H-1 and L-1. In the first case decision-makers are paid \$12 (\$3) if the outcome is an acceptance and \$11 (\$2) if it's a rejection. These treatments will be used as a robustness check of our free-family findings.

In Figure 5 we see the results compared to both their homologous free-rejection treatment and to the N baseline. The most striking feature is that even when rejections can cost 1/3 of the total payoff, the inverted U pattern is still there[11] (i.e. there still are rejections of generous offers!).

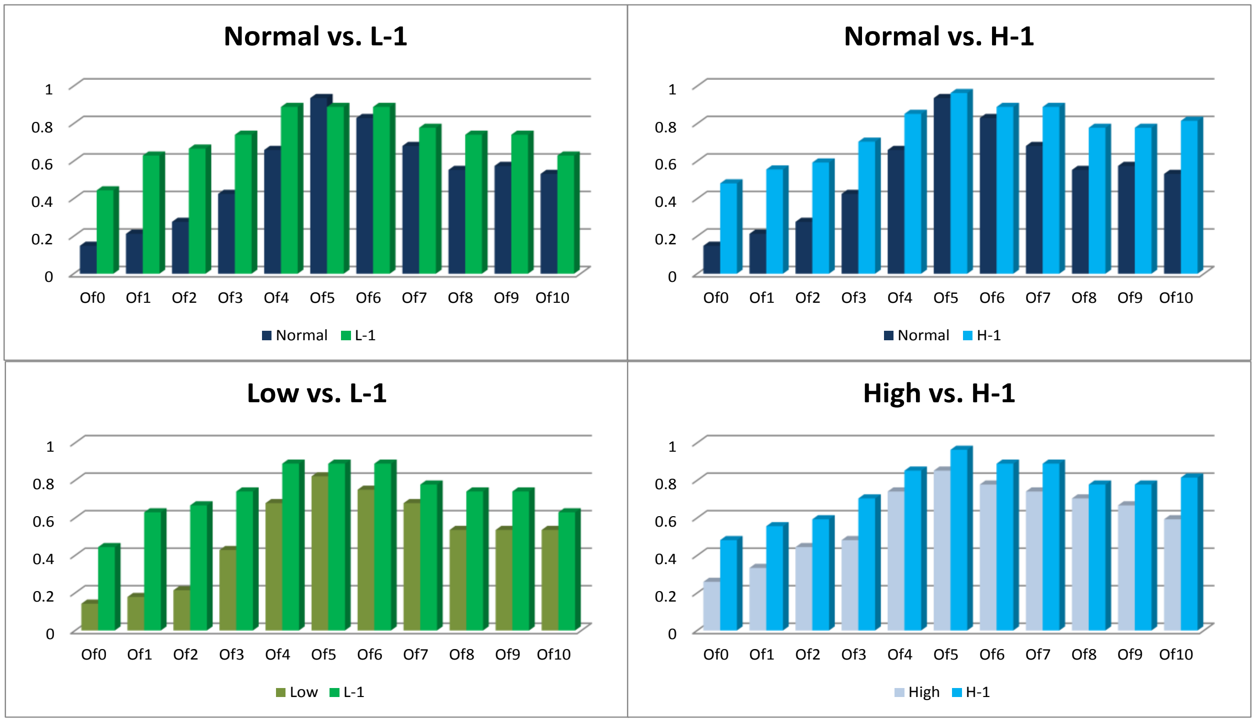

Figure 5: Acceptance rates of L-1 and H-1 vs. Normal, Low and High Treatments

Also noticeable is the similarity between H-1 and L-1 rejection patterns even if the relative cost to reject is so different (Figure 6). As consistent this is with the result to Question 1, it still comes as a surprise. Finally, it also looks like the tails in both costly-treatments are very symmetric; something that would mean that introducing a cost completely eliminates intention-driven rejections.

Question 4: *Do decision-makers act differently under H-1 than under L-1?*

To answer this question we do a Wilcoxon Matched-Pairs Sign-Rank comparing the total number of accepted offers per subject under each treatment, which tells us that there are no differences (p-value=0.6172). Additionally, we run a linear probability model comparing both treatments which shows an insignificant effect of the dummy for treatment (p-value=0.673). A Two-Sided Fisher Test confirms the results (Appendix C).

[11] See Appendix B for Spearman Correlation results

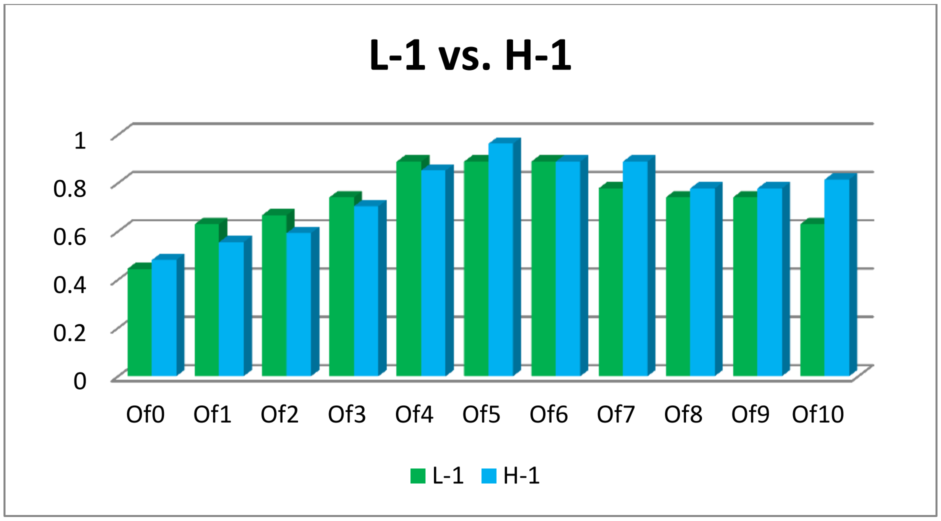

Figure 6: Acceptance Rates of L-1 and H-1

The results presented above seem enough to claim that there is no significant difference between treatments in the costly-rejection family.

Answer 4: *No, decision-makers do not act differently under H-1 than under L-1.*

Given Answer 4, we then wonder if there is any difference between costly and free-rejection acceptance patterns.

Question 5: *Do decision-makers behave differently in costly-rejection treatments than they do in the free-rejection ones?*

We run again a linear probability model, this time comparing the data from L to L-1 and from H to H-1. The results show that not only are L and L-1 significantly different (dummy for L-1 p-value=0.000), but that so are H and H-1 (dummy for H-1 p-value=0.002). To get a clearer idea of where these treatments are different we run a One-Sided Fisher Test[12].

	$0	$1	$2	$3	$4	$5	$6	$7	$8	$9	$10
L-1=L	0.014**	0.01**	0.01**	0.019**	0.058*	0.374	0.163	0.301	0.097*	0.097*	0.333
H-1=H	0.079*	0.085*	0.207	0.083*	0.250	0.175	0.234	0.147	0.379	0.272	0.067*

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 9: One-Sided Fisher p-values Comparing Total Acceptances per Treatment

The results are consistent with our results of the linear probabilistic model, but show an interesting result; most of the differences are found in the LHT of the distributions. In fact looking at Figure 5 we see how the difference across treatments is indeed centered in the LHT, where acceptances in L-1 more than double those in L.

[12] We feel like we can run a One-sided Test because as we can see in Figure 5, acceptances in the costly-treatment are unambiguously higher. A Two-sided Test would not change Low treatment comparisons but would affect H where there would be less significant differences.

<u>Answer 5:</u> *Yes, subjects act differently under costly-rejection treatments than under the free-rejection ones.*

In answering Question 5 we have seen how most of the differences across treatments are in the LHT. Given this result and looking at Figure 5 we then speculate about the symmetry of the tails under costly-rejection treatments.

<u>Question 6:</u> *With the introduction of a $1 cost to reject, are LHT and RHT symmetric?*

To compare the symmetry of these distributions we run a Wald Test (Table 12) comparing the dummies for equidistant offers only within the same treatment (Table 10). The results show a clear symmetry in L-1 but not so much for H-1, where it is a close call except when comparing the two extremes.

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***

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

Table 10: P-values of Wald Test, testing for equality in regression coefficients

On the other hand, when we compare using a Fisher Two-Sided Test, then we cannot reject the null that indeed H-1 is a symmetric distribution except (again) in the very extreme of the tails (Table 11).

	$4 vs. $6	$3 vs. $7	$2 vs. $8	$1 vs. $9	$0 vs. $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**

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

Table 10: P-values of Wald Test, testing for equality in regression coefficients

The results in response to this question give one clear and one ambiguous answer; while the L-1 distribution of acceptances is clearly symmetric, H-1 it is not clear. In the latter, tables 9 and 10 show that the distribution is border-lining symmetry in most points, except for the extremes of RHT and LHT. We take this as a rejection of symmetry for H-1.

<u>Answer 6:</u> in some cases; there is symmetry *under the L-1 treatment, but not under H-1.*

Questions 4 to 6 have helped us analyze the effects of our costly treatment on decision makers. The results seem to imply that under costly-rejection, decision-makers are less keen on showing concern for the proposer's intentions and most rejections are driven by inequality aversion amongst other subjects. We believe that the big increase in acceptances in the LHT and the symmetry (or almost symmetry) of acceptance distributions are proof of this change of mentality.

5. Conclusion:

We have presented a three-player ultimatum game with some results that are at odds with the classic literature on social preferences. The first of these is how the referees of our experiment seem to completely ignore their remuneration when making decisions concerning other subject's payoffs. We believe that these can be interpreted as good news for our legal system as it points towards a potential impartiality of judges when deciding on outcomes that do not affect their payoffs. However, our second result is problematic; under the game structure that we suggest, decision-makers turn into apostles of equality, making a worrisome number of Pareto-inefficient choices.

In the second part of the paper we introduce rejection costs to test the robustness of rejections of selfish offers, the RHT. To our surprise this treatment has its biggest effect in the selfish offers tail, where acceptances more than double in one treatment! Furthermore, the more expensive it gets to reject an offer, the more similar both tails of acceptances get, ending up in a symmetric distribution in our costliest treatment. We interpret these results as decision-makers having a concern for intentions, but these being of second order when compared to inequality between players. So, while rejecting offers is cheap, the intentions of the proposer play a role, but once the price is too high, the concern for intentions disappears.

References:

Bahry, D., & Wilson, R. (2005). “Confusion or Fairness in the Field? Rejections in the Ultimatum Game under the Strategy Method.” *Journal of Economic Behavior & Organization*, Vol. 60 (2006) 37-54.

Brandts, J., & Charness, G. (2000). “Hot versus Cold: Sequential Responses and Preference Stability in Experimental Games.” *Experimental Economics*, 3, 227- 238.

Brandts, J., & Charness, G. (2010). “The Strategy versus the Direct-response Method: A Survey of Experimental Comparisons.” Mimeo.

Bolton, G., & Ockenfels, A. (2000). “ERC:A Theory of Equity, Reciprocity and Competition.” *American Economic Review*, 90, 166-19.

Bowles, S., & Gintis, H. (2002). “Homo Reciprocans.” *Nature*, 415, 125- 12.

Charness,G., & Rabin, M. (2002) “Understanding Social Preferences with Simple Tests.” *The Quarterly Journal of Economics*, Vol. 117, No. 3, pp. 817-869.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007) ‘‘Egalitarian Motives in Humans.’’ *Nature*, 446 (7137): 794-9

Engelmann, D., & Strobel, M. (2004) “Inequality Aversion, Efficiency and Maximin Preferences in Simple Distribution Experiments.” *American Economic Review*, 94 (4), 857-869.

Engelmann, D., & Strobel, M. (2006) "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Reply." *American Economic Review*, 96 (5), 1918-1923.

Falk, A., Fehr, E., & Fischbacher, U. (2005) “Driving Forces Behind Informal Sanctions.” *Econometrica*, 73 (6), 2017-2030.

Falk, A., & Fischbacher, U. (2006) "A theory of Reciprocity." *Games and Economic Behavior*, Elsevier, Vol. 54 (2), 293-315, February.

Falk, A., Fehr, E., & Fischbacher, U. (2008) “Testing Theories of Fairness:Intentions Matter.” *Games and Economic Behavior*, 62 (1), 287-303.

Fehr, E., & Gächter, M. (2002) “Altruistic Punishment in Humans.” *Nature*, 415, 137-140.

Fehr, E., & Schmidt, K.M. (1999) “A Theory of Fairness, Competition, and Cooperation.” *The Quarterly Journal of Economics*, 114 (3), 817-868.

Fehr, E., & Fischbacher, U. (2003) “The Nature of Human Altruism.” *Nature*, 425, 785-791.

Fehr, E., & Rockenbach, B. (2003) “Detrimental Effects of Sanctions on Human Altruism.” *Nature*, 422, 137-140.

Kagel, J., & Wolfe, K. W. (2001) “Tests of Fairness Models Based on Equity Considerations in a Three-Person Ultimatum Game.” *Experimental Economics*, 4 (3), 203-219.

Knez, M, & Colin F., C. (1995) “Social Comparison and Outside Options in Three-Person Ultimatum Games." *Games and Economic Behavior*, 165-94.

Leibbrandt, A., & López-Pérez, R. (2008) “The Envious Punisher: Understanding Third and Second Party Punishment with Ten Simple Games.” Institute for Empirical Economic Research, University of Zurich Working Papers Series ISSN 1424-0459.

Rockenbach, B., & Milinski, M. (2006) “The Efficient Interaction of Indirect Reciprocity and Costly Punishment.” *Nature*, 444, 718-723.

Xiao, E., & Houser, D. (2005) “Emotion Expression in Human Punishment Behavior.” Proceedings of the National Academy of Science, 102 (20), 7398-7401.

Appendix A: Ordering Effects

In order to determine if there are any ordering effects we run a 2 tailed Fisher Test comparing first round treatments against other rounds in the experiment.

P-Values	$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	0.283	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

Table A: Two-Sided Fisher P-values Comparing L-1 and H-1 Treatments.

While most treatments present us with results that show no ordering effects, the H treatment seems to be affected by ordering effects. This can be explained by the lack of observations that we have for first round H treatments (5) compared with third round H treatments (22). This was due to communication problems while running the experiments. If we look at Graph A, it seems like third round subjects behave "normally" while those in the first round H treatment deviate from what we observe in all other rounds.

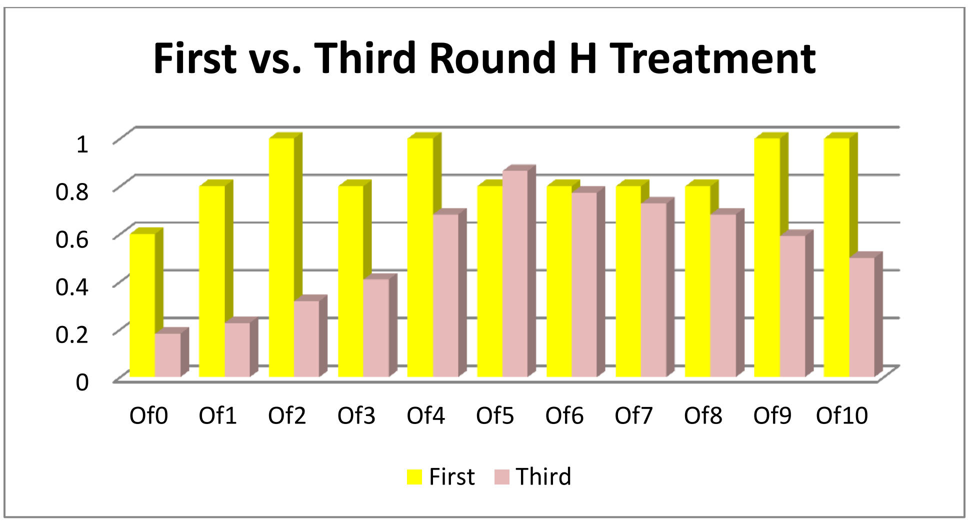

Figure A: Acceptance Rates for H for First (n=5) and Third (n=22) Round

Appendix B:

	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	0.000	0.0012	0.059

Spearman Rank Correlation Results for LHT and RHT of L-1 and N-1 treatments.

Appendix C:

P-Values	$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

Two-Sided Fisher P-values Comparing L-1 and H-1 Treatments.