An Experimental Analysis of the Ultimatum Game:
The Role of Competing Motivations

Loránd Ambrus-Lakatos
Tamás Meszerics

This Draft: 18 June 2001

1. Analysis of the Ultimatum Game

In the Ultimatum Game players are to distribute 100 units among themselves according to the following rules. Player I is the allocator and player II the responder. Player I proposes a division of that sum between the two players. Player II responds by saying yes or no. If the response is yes, the proposed division is implemented. If the response is no, both players get nothing. In game theory, we shall find out what the two players are to do in this situation by applying solution concepts. The Ultimatum Game being a game in extensive form and of perfect information, the appropriate solution concept is subgame-perfect Nash-equilibrium. Accordingly, player I is to propose a 99:1 division; and since one is more than zero, player II is better off accepting this radically unequal distribution of the sum than receiving nothing.

This result is so counterintuitive, that the game has been on the agenda of experimental studies for almost twenty years. In their first report on Ultimatum Game experiments Güth, Schmittberger and Schwarze (1982) found that player I would not infrequently propose a fifty:fifty share, and that she in fact would be unwise to demand more than about two thirds of the sum, since then she would stand about half a chance of being refused. The difference from the subgame-perfect prediction is so robust that it led to a large number of papers, which argue that the observed behavior can best be explained by commonly held fairness norms, instead of the selfish calculus of an "homo economicus" who would care only about pecuniary payoffs (for an overview see Thaler 1988). In these explanations the fairness norm is an exogenous factor that keeps under control, or even erases self-regarding considerations of the players. According to this view, players are closer to something like "homo moralis" when dealing with this particular distribution problem.

Though the observed behavior of the players in the two different roles are both inconsistent with the predictions of game theory, there is some asymmetry in the possible explanations. In fact it is somewhat easier to interpret the actions of the recipients (player II). Refusing a positive, but radically unequal offer may indeed be seen as a clear signal that one cares not just about the monetary payoff but about the distributional characteristics of the allocation as well. In short, refusal of such offers reveals a concern about "fairness". For allocators (player I), on the other hand, there are at least two separate but interrelated motives that could explain the relatively equitable offers. Not only may they care about and evaluate the distributional characteristics in light of a "fairness" norm, they also could simply be concerned with the possibility of rejection from the responders. Although at first sight this latter concern seems to be another manifestation of the same fairness norm, it is safer to state that allocators could in fact be uncertain about the true motives of recipients.

However, there is, evidence that allocators are inclined to choose an equal distribution over a radically unequal one even in the absence of the threat of refusal (Kahneman, Knetsch, and Thaler, 1986), that is when the so-called Dictator Game is played. Further, the same authors found that when subjects were asked about their contingent responses to varying future possible offers, the mean minimum acceptable offer was around 23%. This is then less than what tends to be accepted when the game itself is played, when both the median and the average offer seems to be around 60. Note that this latter experiment asked responders to form a complete contingent strategy, that is to do something else than to respond to an offer actually made. It is natural to suppose therefore, that some sort of fairness consideration enters into the reasoning of each player in the Ultimatum Game.

What this presumed fairness norm is, is not entirely clear however. The most obvious norm that could guide allocators is a distribution of fifty-fifty. This norm may also be derived by applying standard principles of normative collective choice; the Nash-bargaining solution also prescribes equal shares for both players. The norm can even be regarded as focal, as neither the Ultimatum Game nor the abstract bargaining game that may be in the background allow individual characteristics to feature in the problem. Still, while fifty-fifty is far from being a rare outcome in the experimental tests of the game, it is in no way an overwhelming pattern. Again, the average offer tends to be 60.

So the pattern of evidence we need to explain is really this: responders tend to "punish" radically unequal proposals by the seemingly irrational mutually disadvantageous veto, but accept offers that assign a slightly higher share to player I. At least two, somewhat different, interpretations are possible. There might exist a shared social norm, which, instead of pinpointing a unique ratio of "fair" distribution identifies a range of "morally permissible" allocations (say between 50:50 and 70:30). If this is the case, the problem of player I mainly consists in finding out whether her estimation on the lower bound corresponds to that held by the recipient (player II). On the other hand, it may be argued that the asymmetric structural characteristics of the game (the "structural" advantage assigned to player I) in itself generates a particular distributional norm which is no less precise but not identical to the equal share norm. If this is the case, the problem for the allocator is to find out which particular division of the cake conforms to a norm to which both players are willing to subscribe. Note that due to the restricted response opportunity of the recipient it is very difficult to disentangle these two different understandings of the "norm".

Now, positing a norm with vague boundaries poses the difficulty that it can explain just about any outcome. The second possibility opens up another question as well: which structural or institutional characteristics have an effect on the fairness norm? The potentially broad range of structural or institutional variables, that the rules of the Ultimatum Game should comprehensively track, cannot be clustered into a limited taxonomy in order to yield testable propositions. Neither is it easy to discern how these variables could yield norms for distribution. We will return to this issue shortly.

As it is nearly always the case, one feature of the strategic situation stands out as a crucial variable, which is supposed to have an impact on the behavior of subjects: the structure of the information endowments of the players. This has been the center of a large and continuously growing number of experimental analyses of the Ultimatum Game as well. Now, if the game is played only once by two agents, all the relevant information that they could use in choosing what to do is inscribed into the rules. So in order to involve considerations related to information endowments, we need to turn to scenarios where a round of the Ultimatum Game is not played in isolation.

When the Ultimatum Game is played repeatedly, a putatively "fair" allocation, that is one of more or less equal shares, may recurrently obtain and even solidify into a stable multi-period equilibrium. Such repetitive setups introduce the possibility of punishing players who had given selfish offers, or more generally, of reciprocating previous moves. Clearly, fear of retaliation is not the same as a normative conviction about the "right"/"just"/"fair" distribution. Further, it is intuitively obvious that in a repeated game setting with changing roles and changing pairs of players it might be crucial whether the results of the single period games are public knowledge or known to the players of a particular game only. If a social norm is one of the main motivating concerns of the subjects then observing someone making a subjectively "unfair" offer may trigger a retaliatory response in any of the observers, who at a later stage may be paired with the "unfair" person. Such retaliation could be a demand inscribed into the social norm. Then punishment, in turn, can already be anticipated by the allocators in any single period, and change their thinking about the acceptable level of offers. The claim that subjects are more generous to strangers whom they observed to be generous to someone else than whom they saw to be "greedy" in a previous game (in which they did not themselves participate) seems to be valid in a variety of experiments (see Thaler, 1988).

But the suggestion that one ought to involve the idea of reciprocity in attempts of explaining the experimental evidence is equally difficult to accommodate on game-theoretical terms. It would be tempting to claim that the guiding norm of the subjects in these experimental settings is not a perceived ideal distributional share (be it defined by a range or a particular division), but merely a "quasi-behavioral" norm, like reciprocity. Reference to reciprocity seems to be able to explain both the dynamics observed in repeated plays of the game and the acceptable deviation from the "egalitarian" offer. Two considerations, however, make this explanation less convincing. First, the observation, that the dominant pattern of a 60:40 offer emerges in repeated games with changing opponents just the same way as with constant players. This is true even in those cases where outcomes of the individual games remain private knowledge to the players. (Note further that with an innovative experimental design Güth and van Damme (1998) showed that information conditions matter in the one-shot game as well.) Secondly - and perhaps more importantly - the requirement of reciprocity is entirely insensitive to the distributional pattern in the single period game. It only asks that any particular share must be exactly reciprocated and held up over many periods. It is easy to see that any single distributional pattern, if fully reciprocated, yields the same aggregate payoff as the strictly observed fifty-fifty share. Nevertheless, retaliation to an offer that is apprehended as unfair, in the form of rejection, could be regarded as standard motive in the Ultimatum Game. This is because punishment of deviations from perceived norms of justice may indeed be regarded as just.

However, accounting for backward-looking motives such as retaliation can pose a considerable challenge for game-theoretical analysis. It is noteworthy that while one could explain playing 60:40 steadily by a repeated game argument that does not presume any role for fairness norms, this line of argument has not been proposed in the literature. As it is not easy then to account for the evidence of Ultimatum Game by positing the presence of fairness norms, there is a temptation to discard them altogether from explanatory attempts. Indeed, there are prominent experimenters, who attach much less importance to the fairness paradigm. Alvin Roth and Ido Erev (1993) argue that the observed experimental results on the Ultimatum Game are merely a sequence of snapshots of the first steps on an ongoing learning process. They propose a model of reinforcement learning borrowed from psychology, which is consistent with much of the experimental observations but would eventually lead to the subgame perfect equilibrium if allowed to continue sufficiently long. The claim that short- or medium term observations may not be very informative about long term equilibria reached through some learning process is a strong and remarkable contention. Though it contains valuable insights, it ultimately relies on an "as if" argument, in order to vindicate the predictions of game theory. It comes at a price, however. Individual motivations no longer feature in the explanation, since norms by reinforcement learning need not appear in anyone's mind as an action-guiding consideration. Adaptive models may well explain the emergence of norms but by definition cannot shed much light on how these heuristics can and do motivate individuals in their personal decisions. It is our view, however, that in explaining social behavior one needs to refer to reasons for action.

Ken Binmore and Larry Samuelson (1994) have much sympathy with the reinforcement learning argument, but state that it can be harmonized with an explanation that acknowledges social norms of fairness. If norms are seen as endogenous to the process, then applying evolutionary dynamics as the basis of the model may yield predictions that are closer to the observed medium term behavior of experimental subjects. In their understanding, an abstract game may trigger well-established norms that are not particularly adaptable to the central problem of the game. Consequently a new adaptation process starts, and through trial and error learning a new equilibrium-generating norm replaces the old one. However, this view also avails itself to an "as if" argument; in addition, it begs the question of what spans the fairness norm itself.

There should still be a role for learning when the Ultimatum Game is played repeatedly. First, upon the previous arguments, there are good reasons to focus on scenarios when what went on in particular games does not become public knowledge, even if this implies that we abstract from an obvious way in which social learning about the fairness norm itself, and about the fairness of individuals, may be propagated. Next, consider that an equilibrium could arise without any of the players being directly motivated by the aim of implementing a particular distributional norm. There is evidence that when the game is played repeatedly, many allocators who start out with the equal share offer move upwards in later rounds to demand slightly more for themselves. This suggests that there is need for coordinating on a pattern of offers, that is there is a need to reach by mutual adjustment a settlement between the players.

2. The Competing Reasons Account

It is our conviction that players of the Ultimatum Game are concerned by considerations of fairness. But how can one account for these considerations? We forward the thesis that the Ultimatum Game delivers competing reasons for action to the participants. First, as nothing but the order of moves distinguishes the two players, the fair distribution appears to be a 50:50 sharing of the surplus. However the rules also endow the first player with an advantage. As we do not expect people to divest themselves entirely of their aims of getting - if possible - more rather than less, it is natural that player I attempts to take some advantage of her privileged position.. Hence the tendency to offer somewhat more than a half for herself. The boundary between fair and unfair offers is unclear, but 90:10 is clearly beyond it. So a player who makes a 60:40 offer hedges appropriately between the two competing reasons that can ground her action: that of observing norms of fairness and that of pursuing what is in her own interest. However, there is no guidance as to the terms of the trade-off between these two reasons; despite of the importance for player I of settling this, not only because he has a privileged position to affect the outcome, but because of the need to gauge what player II thinks of acceptable hedges. It appears then, that despite of its explicit formal rules the Ultimatum Game, as a distributive problem, lacks a sufficiently succinct structure, at least in the sense that motivations for doing this or that cannot be accounted for in precise terms. It follows that what the acceptable hedge is has to be learned by both players.

According to this account if player I becomes aware of a new feature of the situation that could be regarded as affecting the trade-off between the two competing reasons, the balance between the reasons should shift. First, one can think of a situation in which player I is asked what she would do in the roles of the allocator and that of the responder, respectively. This new feature of the game serves as a perceived additional factor that can guide the resolution of the conflict between the allocator and the responder, as she may then become aware of the possibility to look at the situation from a more or less impersonal point of view. Alternatively, the allocator could receive a chance to make a random draw, the outcome of which determines whether the sum to be shared could increase or not. This added feature could emphasize the importance of her own traits as something that could affect what the desirable outcome could be. If lucky, she may think that it is her who makes the gain for both by actually winning the lottery, and thus should receive a larger share of the resulting surplus. Or she can think that the result of the draw revealed something favorable about her characteristics, namely of being lucky, which in itself could be seen as grounds for "deserving" more of the surplus.

In short, player I in the baseline game grapples with competing reasons, and the indeterminacy could be so vexing that she is prone to squeeze out further reasons from any perturbation of the game, as that can be tied somehow to her actions. Despite the fact that the antecedent lottery is irrelevant to the payoff structure (and, therefore, individual attitudes to risk are also irrelevant) she could interpret luck as an additional reason. Something that signifies "desert" in the broad sense.

Invoking the competing reasons account for the context of the Ultimatum Game may be challenged by arguing that it is only a rephrasing of the “homo moralis” discussed at the outset. This challenge could be sharpened by reference to revealed preferences. That is, there is on the part of both players a concern for both the fairness norm and for pecuniary gains, and the trade-off between these depends on one's preferences. And then in perturbed games, participants may reveal a modified trade-off between the two objectives. However, this view could not explain why after modifying the Ultimatum Game, there would be a change of preferences favoring the motive of pecuniary gains. As the fairness norm is not affected by a perturbation, why would obeying it become less preferred? Also, why does a change in the size of the payoffs affect the force of the pecuniary considerations? In our view, the concern with meeting the fairness norm motivates in a different manner than preferences motivate: it is not a matter of preference whether one keeps to the fairness norm. Accordingly, the competing reasons account avoids the leveling of these two sorts of motivation; thus positing an objective function that, say, expresses a trade-off between pecuniary gains and distance from the fairness norm is to be sanctioned. On this head, we can further argue that our thesis is sufficient to account for the available empirical evidence. Next, suppose on the contrary that such an objective function is invoked in an analysis of our game. We claim that on this supposition the modification cannot change the optimal allocation perceived by the players. In addition, regarding the Ultimatum Game as a Bayesian game skirts absurdity as the possible forms the objective function of player II can take should partly generate the space of his possible types (see the Appendix). Finally, punishment of perceived unfair offers could be part of the dictates issued by the fairness norm. Now the ensuing motivation to reject an exploitative offer is not affected by the perturbation. Rather, the perturbation could only bring about new reasons that enable agents to decide whether the rejection option should be favored.

The Modified Game

Our strategy for testing the competing reasons account employs the following modification of the Ultimatum Game. Player I is to make a random draw before his offer is made. Depending on the outcome of this lottery, the sum to be distributed among the two players may or may not be more than the original 100. Upon the competing reasons account, we expect that those agents who, who win on this lottery - in the sense of gaining a larger sum to be distributed - might tend to propose themselves more in relative terms. (That is: if they proposed, say, 55:45 in the baseline Ultimatum Game, and the draw increased the sum by a factor of four, they might make in the modified scenario a 250:150 offer.) Again, they may behave as if they think they deserve more after a lucky draw.

This, however would be an anomaly, as upon standard game-theoretical solution concepts such a proportional increase in the surplus should not affect the relative shares to be implemented. The argument for the 99:1 offer is independent of the overall amount featured in the game. On the other hand, the perceptions of what fair shares are, should be equally insensitive to the level of surplus to be distributed. Indeed, Cameron (1995) reports that size of the surplus does not affect outcomes in the Ultimatum Game. Furthermore, it is common to conceive of these fair shares as "immune" to that sort of luck that perturbs the baseline game. One may even state that this "luck" is no different from the "luck" that places particular persons in the role of player I. Many go even further and claim that fairness in distribution indeed demands purification from all effects of luck.

Hoffman and Spitzer (1985) already explored some aspects of the effect luck may have on the choices of subjects in bargaining situations. In their experiments allocators were asked to choose between an arrangement that gave all of $12 to the allocator and nothing to the recipients, or an outcome where $14 had to be divided, but the division had to be agreed upon between the two players. (Not identical, but very close to the Ultimatum Game.) The overwhelming majority of subjects chose the second option with an equal division of $7 going to each party. Next, the authors separated four treatment conditions: allocators chosen by the flip of a coin and told that they are designated as allocators; allocators chosen by the flip of a coin and told that they earned the right to be allocators; allocators determined by winning a simple game of NIM and told that they are designated as allocators; allocators determined by winning a simple game of NIM and told that they earned the right to be allocators. Hoffman and Spitzer found that moving from simple luck to luck coupled with some effort (winning the game) had no significant effect on the distribution when allocators are told that they are designated. Those subjects, however, who were told to have earned the right to be allocators (after the flip of a coin or winning the simple game) kept significantly larger sums for themselves.

Partly on the basis of these experiments, Frey and Bohnet (1995) offer a classification of property rights according to the extent that they induce fairness norms. They claim that "undefined property rights" ask for an equal share in a distributive situation, while property rights defined by "luck", "by a gift", and "earned property rights" suggest underlying fairness norms that allow for a successively increasing share to be kept by the allocator in the same situation. Although the taxonomy of particular institutional variables called "property rights" is an innovative one, we think that identifying fairness norms through this concept is at least a questionable one. First of all the baseline case of undefined property rights seems to be dubious. The authors` hypothetical example of player I finding an object on a deserted beach can be just as easily classified as an operation of luck. But beyond the exact scope of the respective categories we think that the introduction of the concept of property rights as a particular type of institutional features opens at least as many questions that it seems to solve. Even if we disregard the normative problem of exactly what personal prerogatives constitute that realm for which we use the shorthand of "property rights", it is not obvious how these rights relate to norms of fairness in distributive situations. Also, it seems to us less than convincing that if one finds a divisible object on a deserted beach then one has an obligation to share it equally with the first appearing stranger disregarding any and all of the personal attributes of that particular stranger. On the other hand we also find it problematic that "property rights" seen as institutional feature can be established by experimenters through verbally stating that the allocator had "earned" that right. It may be dangerously close to inducing a norm through the use of the inherent authority of the experimenter, which seems to be closer to some kind of framing effect.

Finally, institutional features that are to affect how the game is played ought to be represented by the rules of that game. However, the Ultimatum Game is completely defined by its rules, and the only rights players have is to play the game. The chance to gain some money is given to them by the experimenter, their task is to distribute (or to allocate) a certain sum. As the rules are given, even those who are tempted to bargain have only these rules at their disposal. Again, anything more there may be to the situation can only be conveyed by payoff-irrelevant features, that of course could include communication from the experimenter.

There are experimental results that lend support to our thesis in a more indirect way. In an experiment of Tversky and Shafir (1992), participants were first shown cards with prizes written on each. Then these cards were put into a deck and two of them were handed over to the decision-makers. They could have either accepted one of them, or ask for a third one. This latter option was costly. In one scenario, the two prizes were A and B and neither dominated the other. In the other scenario, the two prizes were A and C, where A clearly dominated the other. Nevertheless, a significantly higher percentage of the participants asked for a third card in the first scenario. This happened, supposedly, because they did not have enough reasons to choose the very same A they tended to choose in the other scenario. Rubinstein (1999) reports an other experiment where 80 percent of the participants deferred her favorite action in the Battle of the Sexes game upon learning that the other already made his move. Clearly, in the Battle of the Sexes game there is no decisive reason to choose either of the two available actions. However, having an access to information that is regarded as irrelevant from the game-theoretical point of view apparently gave, in Rubinstein's sample, a further reason to defer one's choice to the other.

3. Evidence

The Experiments

Four experiments were held at the Central European University between the spring of 1998 and the fall of 2000. In each of the first three experiments 6 MA students were selected from applicants who responded to an advertisement on a "social science experiment". Subjects were led into a classroom and a number (1 to 6) was assigned to them randomly. In each case, after reading and discussing the instructions the first session started. Players were matched in a truncated tournament fashion and each played six times the standard Ultimatum Game where they had to distribute HUF 500 (Players 1, 2, 3 played against 4, 5, and 6 first as allocators then changing roles). The offers and the results were communicated through the experimenter and therefore were private information. No pair of players was aware of the results of the other pairs.

In the second sessions of these experiments the perturbed game was played in the same fashion. Before making an offer, allocators rolled a dice and the results determined the sum they had to allocate (between HUF 500 and HUF 1,000). With probability one-sixth they gained nothing, with probability one-third they gained an additional 100 HUF, with probability one-third the additional gain was 300 HUF, and, finally, with probability one-sixth an extra 500 HUF was received. The results of rolling the dice were public knowledge, while the other information conditions remained unchanged. In each game in both sessions players were asked to evaluate their offer (or the offer they received) by indicating on their game-sheet whether the offer was fair and whether it was acceptable. These evaluations were not communicated either to the opponents or other players.

In the fourth experiment the 8 subjects were undergraduate economics students, who played two games each in three separate sessions. Opponents in each game were matched randomly.

The first two sessions had very similar conditions as in the previous experiments. In this experiment the lottery in the second session yielded a sum of either 200, or 500, 800, 1000, 1500, and 2000 HUF, each to occur with equal probability. In the third session, instead of rolling the dice, players in the role of allocators had to answer a trivia question which determined the total sum that was to be divided (a right answer yielded HUF 1,500, while a wrong one HUF 500)

The Hypotheses

Beyond the competing reasons account initially we formulated four rival hypotheses, which the experiments were expected to test. Though each of them are on a somewhat different level of abstraction, all three were plausible a priori and the experimental results were hoped to prove or falsify them, or at least to discriminate between their comparative strength.

Total transparency of payoff irrelevance. This hypothesis would claim that the structure of the perturbed game is still fully transparent to the players. They immediately understand that the antecedent random draw is irrelevant to the distribution problem. The prediction of this hypothesis is: no effect on relative offers (second session games continue with the same proposal ratios observable at the end of the first session games).

General perturbation effect. It also seemed possible that the perturbation of the baseline game simply confuses the participants in their choice of strategy. The introduction of a payoff irrelevant action would induce divergent (and maybe individually inconsistent) behavior in the participants. In one sense the effect is similar to a sudden increase in noise level. Prediction: offer ratios in the second session would deviate from the mean offer ratio established in the first session, but deviations would be non-patterned (random in the limiting case).

"Luck as strong entitlement". This hypothesis is based on the argument that participants perceive the surplus generated by the antecedent random draw as a direct entitlement resulting from their physical action of rolling the dice. It is essentially a stylized and at the same time extreme formulation of the well-known Lockean theory of primary acquisition. This hypothesis would predict that in second session games the proposer retains the entire surplus resulting from the draw and in absolute terms the proposed share to the responder remains constant across the two sessions.

"Competing reasons". According to this hypothesis, as the sum that can be distributed in the second session increases, the allocators will claim more for themselves than in the baseline scenario. But they do not attempt to claim all the surplus that is yielded by a lucky draw. As a general observation we should note in advance that experimental tests suggest that the observed high offer ratios are insensitive to the magnitude of the payoff. The work of Cameron (1995), mentioned already above, reports the most convincing experiment in this respect. With this in mind we could exclude those accounts that use the magnitude-sensitivity argument. However, we leave room for a size-effect that results from a lucky draw.

"Establishing the norm". In our experiments, the same subjects were first confronted with the standard Ultimatum Game and, in the second sessions, with the modified game as well.

Therefore they are exposed to a repeated play of versions of the same game. We can reasonably expect therefore that they will mutually establish some distributional norm, they will learn in the successive sessions what offers are reasonable to make. This hypothesis is compatible with the competing reasons account. However, taken in itself, it does not predict any direction in which the offers made in the second sessions are to change. Note that the competing reasons account could have been tested by using two separate groups for playing the standard and the modified game, respectively. In this case, the learning effect covered by this hypothesis could not arise and thus the test would, supposedly, yield more clear-cut results. However, with our experimental design we can compare the behavior of the very same individuals, across sessions.

The predictions of the rival hypotheses can be summarized in a simple table:

Table 1: Hypotheses and the observations they are to yield

<table>
<tr><th colspan="2">Observation</th><th>Hypothesis</th></tr>
<tr><td colspan="2">I. No change of percentage offers between sessions</td><td>“Total transparency of structure”</td></tr>
<tr><td rowspan="4">II. Change of percentage offers between sessions</td><td>II/a Random change</td><td>“General perturbation effect”</td></tr>
<tr><td>II/b Entire surplus from random retained by allocator</td><td>“Strong entitlement”</td></tr>
<tr><td>II/c Moderate decrease in Percentage offer</td><td>“Competing reasons”</td></tr>
<tr><td>II/d Gradual convergence To a mutually acceptable Offer</td><td>“Establishing the norm”</td></tr>
</table>

The Results

In the first three experiments, 18 subjects played a total of 54 standard and 54 modified Ultimatum Games. (The detailed results of all the four experiments, in form of summary tables, are available from us on request.) From Table 2 below, it is clear that between the first and second sessions the results show a moderate but still clearly observable change in the proposal ratios. In two experiments the median offer to recipients decreased while in one experiment the median remained the 50% offer. The median offer for all standard games was 45%, while the median for all perturbed games was 40%. The mean value was 38.9% and 36.8%, respectively. In the four experiments taken together (140 games), the average offer for the standard game was 38.6 and in the modified game 37.7. We consider that this fact is significant enough to allow us to discard the total transparency hypothesis. Indeed a large number of the participants perceive the change in the structure of the game, and it has an effect on their choice of action.

On the same account it is also clear that this non-dramatic, but observable change shows a vague pattern. The mean offer ratio slightly increased in all but one experiment. Though - as usual - there is significant variation across individual participants, it is still not a random change. Therefore the hypothesis on the overriding strength of the general perturbation effect can also be discarded. This is not to say, however, that the results would show that all participants perceived and interpreted the change in the structure of the game in a uniform way.

For an evaluation of the competing reasons account, we calculated the offers in the modified game depending on the additional sum that the random draw yielded. The results are shown in Table 2.

Table 2: Average and median offers in the first three experiments (numbers in parentheses indicate sums in HUF)

Scenario	Mean offer	Median offer
Standard Game (500)	38.9	45
Modified Game (500 plus)		
No or minimal increase (0, 100)	40.9	41.6
Substantial increase (300, 500)	33.9	37.5
Total for all Modified Games	36.8	40

(Note: mean and median offers are normalized to percentages.)

Clearly, the larger the sum that was gained by the random draw the less the allocator offered to his or her partner. We note first that in three out of the small amount of cases (9) when the subject won the maximum extra amount of 500, those who were thus favored were the only two subjects, out of the 26 in all of our experiments, who never offered anything else than the fifty-fifty shares. And we also note also that for those who made no wins, the average offer increased to 46.9, and the median offer 50.

The first two sessions of the fourth experiment delivered a further 32 observations. The results are summarized in the next table.

Table 3: Average and median offers in the fourth experiment (numbers in parentheses indicate sums in HUF)

Scenario	First Round		Total	
	Mean	Median	Mean	Median
Standard Game (1000)	37.5	40	37.5	40
Modified Game				
Decreased or same sum (200, 500, 800, 1000)	41.0	40	42.0	40
Increased sum (1500, 2000)	32.5	32.5	37.5	40
Total for Modified Games	37.8	40	40.6	40

(Note: mean and median offers are normalized to percentages.)

It appears that those allocators who actually lost money in the draw are more lenient to accommodate their counterparts than those who did not lose or even won some additional surplus. Those who affected a considerable loss in the first round offered 43.3 on average. Only in five out of sixteen cases did subjects make wins over the baseline amount; three of these cases occurred in the first round of the second session. In them, the average offer indeed decreased substantially. However, all of the these three offers were rejected, and the only subject who won in both rounds offered, accordingly, 46.6 to his or her partner in the second round. Note that due to the problem of small numbers, we need to practice caution in the interpreting the results of this fourth experiment.

We conclude, with one caveat that we will discuss at the very end of the paper, that the competing reasons hypothesis should be confirmed.

As far as the strong entitlement theory is concerned, we can safely argue that the results do not conform to its predictions. The number of cases where the offer ratio is identical or even close to the "surplus retaining offer" is small and even in those cases it is not sustained across games. The relationship between luck and entitlement seems to be much less direct then it is understood by the stylized Lockean theory. This is also confirmed by the results of the third session of the fourth experiments. Here, again, an allocator who could have answer a trivia question could have made an offer about splitting 1500 HUF, if he or she failed this was 500 HUF. Whereas in the standard game (played in the first session of this experiment) the mean offer was 37.5, in this second type of modification the mean increased to 45 (the median also increased, to 45.3). It seems that the modest merit of having answered successfully a question not only did not create an entitlement to the enlarged sum, but established a bond between the subjects. This appears to be the case even if we consider that the “learning effect”, to be discussed shortly, could have been in full force in the third session of our last experiment. Note, however, that the mean offer of those who failed to give a correct answer was 46.3, while for those who managed to cope with the question was 44. Obviously a caveat - similar to the one stated above - needs to be added: our results do not exclude that in at least some games a reason similar to the strong entitlement hypothesis could not have motivated the proposer.

From the 156 games only 19 (12.2%) resulted in rejection. The median rejected offer was 20%, the mean 22,1% and the spread was between 0 and 40%. In itself the low number of rejections does not allow for a straightforward interpretation. In one sense it is an indicator of success, since the players secured a large overall payoff. Beyond this observation we cannot say whether this success also means that allocators were able to gauge correctly the shared distributional norm, or they were overcautious in their offers. At the same time, the analysis of rejected offers sheds some light on the complexities of whatever learning process may be at work. One of the most plausible assumptions about the dynamics of the game says that when a player experiences refusal she will increase the offer in subsequent games. There were 6 players, whose offer (at least once) invited refusal in the standard game. Four of them duly increased their average offer in the perturbed game. All in all 10 players increased their average offer between the standard and the perturbed games, 3 offered the same amount on average and 13 decreased their average offer. This balanced individual-level result does not seem to support our hypothesis. It would not be unreasonable, however, to say, that only 4 of the average increases should be compared to the 13 decreases, since 4 of the 10 cases may be explained -so it seems to us - by the prior experience of refusal. Also, the three players who did not change their average offers across sessions include the two who never offered anything but the fifty-

fifty share to their partners. Table 4 provides a summary of these figures. We should note that this motivational argument works on a different dimension than the possible effect of the magnitude of the extra money won through lottery. Due to the small number of observations (19 rejections) controlling for both dimensions would not yield enough cases in the possible types to make a substantial difference in the present argument.

Table 4: Number of players who decreased or increased, respectively, their average offers in the modified game

	Experienced Rejection in the First session	Did not experience Rejection in the First session	Total
Number of players who decreased their average offers in the modified game	2	11	13
Number of players who increased their average offers in the modified game	4	6	10

(Note: average offers are normalized to percentages.)

On the other hand the responses to the fairness/acceptability question show that recipients were also reluctant to use the self-harming veto weapon in all cases when they considered the offer unfair. In 26 out of the 156 games (16.6%) recipients indicated that they thought the offer was unfair but acceptable. Since this evaluation was not communicated to allocators, it could not have been used strategically to influence the beliefs and behavior of the other players. We can safely conclude that the range of fair offers is smaller than that of the acceptable ones, which has significance for the interpretation of other experimental results on distributive problems. Only one player seemed to display a strong egalitarian understanding of the fairness norm and considered as unfair every offer below a 50% share even in the role of allocator. Significantly, this evaluation did not preclude the player from offering a less than equal share in the perturbed game. On the other hand none of those players, who evaluated some offers as "unfair but acceptable" in the role of the recipient reported the same evaluation when making exactly the same offer to someone else. There seems to be a strong tendency for subjects to consider all of their own offers as fair, which indirectly shows how much importance subjects attached to this normative criterion. Whether their inconsistent attitude to fairness was a case of dissimulation or self-delusion cannot be determined from the data. It needs to be mentioned that the largest rejected offer (40%) was the only case when the recipient thought the offer fair but still unacceptable.

There is at least one rival account of the results that at this point we are unable to exclude. It can be argued that the length of the sessions (and therefore the introduction of the perturbation) was exogenously determined. It did not take into account the dynamics of medium term trial and error learning. Therefore it is possible, that the results that we interpret as showing the effect of competing claims can also be understood as confirming the “Establishing the Norm” hypothesis. That is, at this moment the results can also be seen as consistent with some form of "learning effect spillover" that is at work between sessions. We would like to repeat, however, that the two accounts are not necessarily exclusive. In fact the competing reasons description may be seen as providing the initial dynamics that can be reinforced by an unspecified learning process.

References

Binmore, Ken, and Larry Samuelson, "An Economist`s Perspective on the Evolution of Norms" *Journal of Institutional and Theoretical Economics*, 1994, 45-63.

Cameron, Lisa, "Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia" Working Paper, 1995, Princeton University.

Frey, Bruno S., and Iris Bohnet, "Institutions Affect Fairness: Experimental Investigations" *Journal of Institutional and Theoretical Economics*, 1995, 286-303.

Güth, Werner, Eric van Damme, "Information, Strategic Behavior, and Fairness in Ultimatum Bargaining: An Experimental Study" *Journal of Mathematical Psychology*, 1998, 227-47.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze, "An Experimental Analysis of Ultimatum Bargaining" *Journal of Economic Behavior and Organization*, 1982, 367-88.

Hoffman, Elizabeth, and Matthew L. Spitzer, "Entitlements, Rights and Fairness: An Experimental Examination of Subjects` Concepts of Distributive Justice" *Journal of Legal Studies*, 1985, 259-97.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler, "Fairness and the Assumptions of Economics" *Journal of Business*, 1986, 285-300.

Roth, Alvin, and Ido Erev, "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Medium Term" Working Paper, 1993, University of Pittsburgh.

Rubinstein, Ariel, “Perfect Equilibrium in a Bargaining Model”, *Econometrica*, 1982, 50, 97-109.

Rubinstein, Ariel, “Lessons from a Course in Game Theory”, Working Paper, 1999, Tel-Aviv University.

Thaler, Richard H., "Anomalies: The Ultimatum Game" *Journal of Economic Perspectives*, 1988, 195-206.

Tversky, Amos and Eldar Shafir, “Choice under Conflict: The Dynamics of the Deferred Decision”, *Psychological Science*, 1992, 358-361.

Appendix:
A Formal Analysis of the Ultimatum Game

1. In this Appendix, we will provide a technical summary of some of the arguments that were considered in the main text. The Ultimatum Game is an extensive form game of perfect information. It is played by two players, I and II, so N={I, II}. They are to reach an agreement about sharing a certain sum (specified in monetary terms), the level of which is V. This they have to do according to the following rules. First player I gives an offer of the form (x, V-x) to player II, who in turn may either accept it (A) or refuse it (R). We further stipulate that x has to be a multiple of a certain unit, $u>0$. If an offer is accepted then the payoff of player I is x and the payoff of player II is V-x. If an offer is refused then both get 0. So the set of histories in this game is $H=\{(X,(A,R)): \forall x\in X,$ $x=uk$, where $k=0,1,\ldots,V/u.\}$. Finally, we assume first that both players simply prefer more money to less. Further, after an offer of (V,0) the second player always prefers R to A.

Before we proceed, we lay down the conventions that offers are presented in ordinary parentheses; strategy profiles in curly brackets; and allocations or outcomes in square brackets.

Clearly, the set of strategies of player I can be identified as the set X. For player II, the set of strategies is comprised by functions that assign A or R to each offer V-x. It is reasonable to stipulate that these strategies are monotonic in the following sense: they take the form of "accept any offer more or equal than a certain offer you receive, and refuse each offer below that level". So the possible strategies of II can be conceived of as a step-function $\alpha(x)$, that assigns the value 1 to those offers that are to be accepted, and the value 0 to those offers that are to be rejected. Given monotonicity, we can use the shorthand α for any strategy, where, of course, α stands for the smallest offer player II still accepts. By assumption, we must have $\alpha>0$.

The set of Nash-equilibria in this game is easy to characterize. First, $x=V$ cannot be part of a Nash-equilibrium, as $\alpha>0$. Denote candidate equilibria by the ordered pair $\{x;\alpha\}$. Any such pair for which $x<V$ and $\alpha=x$ is a Nash-equilibrium, as x is a best response to $\alpha=x$; and also player II cannot do better by choosing $\alpha=y$ such that $y\neq x$. This is because if $y>x$, he will then get 0; and, on the other hand, if $y<x$, he will not do better than by choosing $\alpha=x$.

There is, however, only one subgame-perfect Nash-equilibrium (SPNE), namely, $\{V\text{-}u;V\text{-}u\}$. If player II observes any offer smaller than V, it will be preferable for him to choose A. So in the SPNE outcome, player I should offer u to her partner and he accepts that.

Note that these results hold irrespective of the level of V. Further, it goes without saying that payoff-irrelevant actions will not affect the results, either.

2. The unique subgame-perfect Nash-equilibrium yields a grossly uneven distribution of eventual payoffs. What shall the two players do if they regard the Ultimatum Game as a scenario in which a collective choice problem has to be solved? Given that the range of possible outcomes of the

Ultimatum Game is identifiable, one can map it into a cooperative bargaining problem. For this, the bargaining set can be reasonably taken as $X(V)=\{[x,y]\in R\times R: x\geq 0, y\geq 0, \text{ and } x+y\leq V \text{ for some } V>0\}$. The disagreement point is, of course, $D=[0,0]$.

It is sensible to suggest the Nash-solution should be used in this associated cooperative game. The Nash-solution here is $[V/2,V/2]$. This outcome may be attained if in the Ultimatum Game the strategy pair $\{V/2,V/2\}$ is played; however, as we have seen, this profile yields a Nash-equilibrium, but not a subgame-perfect Nash-equilibrium. It goes without saying that the Nash-solution could be implemented by non-cooperative games other than the Ultimatum Game; where these non-cooperative games ought to have the same associated cooperative bargaining game as the Ultimatum Game has.

3. Experiments with the Ultimatum Game show that the SPNE outcome is almost never observed. Two explanations are usually offered to account for the empirical evidence. Both respect the SPNE as the most reasonable solution concept for any extensive game of perfect information.

According to the first explanation, the second player's preferences are misspecified in the above description of the game; it is unlikely that someone cares only about pecuniary payoffs in a setting like the Ultimatum Game. Rather, most persons are concerned with the equity or fairness of the eventual outcome as well. That is, they dislike outcomes that may be far away from what can be taken as the fair one, that is, the one induced by offering $V/2$ and its subsequent acceptance.

Suppose then that someone in the role of player II has "fairness costs" in the following sense: he evaluates eventual outcomes by the payoff function $F(z)=S-x-\max(0,x-z)$. Payoffs after rejection remain the same. Here z could be interpreted as a benchmark. Below this benchmark the player in question will not experience resentment about the fairness of the offer of the other; but if $x>z$ he will experience linear resentment costs at the level $x-z$. As an example, we could take $z=0.6$. Then if the offer happens to be $x=0.7$, this player II will get overall (pecuniary plus fairness) payoffs of 0.2 if he accepts that offer. Further if, $z=0.5$ and $x=0.8$, then the overall payoffs for the second player in case of acceptance will become –0.1, that is less than that from refusing the offer.

If we take anyone in the role of player II having these sorts of preferences, then the set of SPNE will change accordingly. Given z, he will only accept an offer x for which $V-x>\max(0,x-z)$, that is for which $F(z)>0$. So, $\alpha(x,z)$ prescribes acceptance if $x\geq(V+z)/2$ and it prescribes rejection if $x<(V+z)/2$. Therefore, the SPNE strategies are $\{(V+z)/2,(V+z)/2\}$ in this case. There is then again a unique subgame-perfect Nash-equilibrium.

We have to notice, however, that z is not observable to player I. If it was, she, in the current case, should give an offer of $(V+z)/2$. Given that it is not, it is left to her to judge what the resentment benchmark for her partner is. Nothing guarantees that each and every person has the same level of z, that is the number characterizing the attitude to fair outcomes in this game. Consider that we can take this revised formulation of the Ultimatum Game as a Bayesian game where potential values for z identify types ("fairness types" in our case). Then all we can suggest to player I is to assess her individual beliefs concerning which of the various possible types she is confronted with. If we suppose that $z\geq V/2$, this leaves her with $(V/2u)+1$ potential sorts of partners. Upon listing the potential types her partner may have, player I has to proceed to calculate x as the offer that maximizes expected gains given her beliefs. But it is difficult to accept that ordinary players use such a calculus for determining their optimal strategies.

On the other hand, one may argue that there is a focal value for z. There is only one obvious candidate for this value, namely, V/2. Then we first note that it is rare that players in experiments reach [3V/4,3V/4], the requisite SPNE outcome. Next, it could be argued that the assumption of linear resentment costs is arbitrary, while the suggested [3V/4,3V/4] equilibrium outcome depends on this assumption. This argument is correct. But we note that any assumption concerning the form which resentment costs may take is equally arbitrary, as it is not specified by the rules of the game. Then if player I is further uncertain about the exact functional form in which resentment is expressed, she faces even more Bayesian types than in cases in which she may be sure that the resentment costs are as described above. We conclude that the suggestion that players play the Ultimatum Game as if it was a Bayesian game is beset with grave difficulties.

4. According to a second suggestion for revised analysis, it is a mistake to consider only one actual play of our game, in isolation. Rather, we have to assume that players see themselves as playing it in a potentially repeated setting. Then the Ultimatum Game as described above is the stage-game in a corresponding repeated game. By standard arguments, it can be assumed that the stage game is thought to be repeated an infinite amount of times. Now, the explanation continues, players may force each other to keep to some more or less fair way of playing by using punishment strategies. Accordingly, one does not need to posit that players have concerns, that is payoffs, that are not specified explicitly by the scenario of the Ultimatum Game.

Equilibrium strategies in the repeated game could take the following form. Each player starts with playing his or her part in a cooperative strategy profile, which is {V-x', V-x'}. Naturally, we are interested in cases when x' is substantially larger than u. If player I deviated from her strategy in the profile, they are to revert for T periods to the first punishment strategy pair, that of playing {0,V}; once this is over they move back to the cooperative phase. If, on the other hand, player II deviated in the cooperative phase, they are to revert to the second punishment strategy pair, prescribing the play of {V,0} for T' periods; once this is over they move back into the cooperative profile. Further, if player I deviated from the prescribed strategy in the first kind of punishment phase, then they start to count the T periods again; if she deviated in the second punishment phase then they move back to the first punishment phase. Similarly, if player II deviated in the second sort of punishment phase they start to count the T' periods again; and if he deviated in the first punishment phase then they proceed to the second one. Finally, assume that a potential deviation of player I has priority over the deviation of player II.

Suppose next that both players use the same discount factor $0<\delta<1$ for evaluating payoffs across periods. We submit, without further argument, that there is a parameter constellation for the sixtuplet $<\delta, T, T', V, u, x'>$ for which playing {V-x', V-x'} always is a subgame-perfect Nash-equilibrium of the repeated game. It follows that the fair outcome [V/2, V/2] can be implemented by using these strategies.

We recall that there is a similar conclusion for the analysis of the infinitely repeated Prisoner's Dilemma game: provided that the discount factor of the players is high enough, attaining the symmetric Pareto-optimal outcome in each period can be implemented by using punishment strategies. But what could x' be in the repeated Ultimatum Game? Note that the level of x' has to be specified in order to identify the trigger of resorting to punishment. If indeed the repeated game analysis is offered as an explanation of the experimental evidence, it has to be accepted that x' cannot simply be taken as V/2, as this is not what the evidence shows. Short of that, there seems to be no other focal value for x'. This means that the Prisoner's Dilemma game does not offer an altogether close analogy, as it has a clear focal point of the appropriate sort, namely, playing {No Confess, No Confess} until anyone would have deviated. Also, we cannot simply say that players can coordinate on some value for x', as there is absolutely no agreement, in the Ultimatum Game, between the two players concerning the ranking of the potential outcomes. It is a well-known difficulty for repeated games arguments that they cannot in a principled way exclude the possibility of multiple equilibria. To this one has to add, in the case of Ultimatum Game, the problem of the absence of any focal equilibrium.

5. The focal equilibrium problem may be confronted by making a further suggestion. According to this, a proper repeated game analysis should use the idea of reciprocity. Here players are supposed to learn that they should avoid an exploitative offer in the role of player I because they would not accept that in the role of player II. Simple reciprocity is to be regarded as the focal behavior in the Ultimatum Game. For the sake of completing this argument, one may conceive of a random matching arrangement in which there are many players, in which the behavior of individuals is public knowledge, and in which people alternate in the roles of player I and II, respectively. Restricting the analysis to cases in which only what happened in the preceding period is remembered, the reciprocity argument suggests the strategy of always offering to a player II what he or she did offer in the preceding period. But the very same suggestion can be made in confines of scenarios that have only two players, who alternate in assuming the two respective roles. In formal terms, one may wish to cash out this proposal by specifying Tit-for-Tat-like strategies to be used by the players.

Unfortunately, this argument is question-begging. This is because players could just as well coordinate on always giving V-u in the role of player I and accepting that in the role of player II; in the long-run this would bring (almost) the same payoff as always giving V/2 offers in the role of player I and always accepting that in the role of player II. This is because in the Ultimatum Game any accepted offer yields the same total payoff for the two players.

Indeed, one does not observe in experiments a pattern in which people always give V-u offers and this is always accepted. Rather, what seems to be appropriate to do is to give more or less equitable offers on each occasion one can do so. This is what needs to be explained.