

# The Ultimatum Game Revisited

Jan Tullberg

*SSE/EFI Working papers series in Business Administration  
No 1999: 2 April 1999,  
Revision January, 2002*

## **Abstract**

The starting point of this article is the result of one ultimatum game experiment - one of many showing a huge deviation from the predictions of micro theory. However, further analysis gives an explanation of subject behavior that deserves to be seen as rational, if assumptions, such as a total secrecy resulting in no effects on reputation, are questioned. Responders' actual behavior can be understood as adjusted to generally realistic reputation effects, and the choices of the Proposers are surprisingly attuned to actual Responder demands. If seen in this light, the subjects seem to understand the situation and behave accordingly.

Keywords: Ultimatum, buffer, reputation, spite, altruism  
JEL classification code: C91, M14

Jan Tullberg (jan.tullberg@hhs.se)  
Stockholm School of Economics  
Box 6501  
SE - 113 83 Stockholm  
Sweden

# The Ultimatum Game Revisited

Jan Tullberg

A question that has interested several researchers is the influence of economics and rationality on cooperation. According to micro theory as used in game theory, cooperation is hard to obtain and many theoretical analyses end up as seeing defection in the Prisoner's Dilemma as the rational result. This is not the behavior obtained in experiments or observations, but even the critics of micro theory seem to consider it partially right; they also believe that if people behaved rationally they would not cooperate. Implicitly or explicitly, the hope therefore goes to unselfish norms - instead of reflection - to block rampant free riding and make people work together.

I disagree with both groups and hold that cooperation can be most rational and there are possibilities to find a satisfactory explanation of what is going on in experiments and "out there". My empirical base will be the results of an experiment conducted by Carter & Irons (1991). I will use that material, and also more detailed data that they have had the generosity to provide me with. The paper aims at showing that their results make sense in a model that is pro-rational, even if not in line with micro theory.

Carter & Iron use experiments with an ultimatum bargaining game to investigate whether economists differ from others in their behavior and, if so, test two hypotheses as steps to better understanding, (this design inspired by Marwell & Ames 1981). The "learning hypothesis" proposes that studies of economic theory influence economists to see things in a different way, while the "selection hypothesis" propose that a difference in attitudes comes first, and is a reason for students to choose economics instead of other walks of life.

## 1 The experiment by Carter & Irons

For the purpose of investigating these two hypotheses Carter and Irons used four different subgroups, obtained by splitting the subjects in accordance with two variables. One was of course economists and non-economists and the other was freshmen and senior students. A comparison between freshmen economists and freshmen non-economists would indicate selection effects, and a comparison between freshmen economists and senior economists would be a component of the learning effect. Since it is the learning of economic theory that is of interest, the difference between freshmen non-economists and senior non-economists can be used to separate general learning and maturing effects with no connection to economic theory.

The experiment itself had several stages. In a first stage the subjects were asked to make as many words as possible out of some given letters. They were then told that they later would be randomly paired with another person for a second game. The person with the higher score in the word game would get the role of "Proposer" and the one with the lower score the role of "Responder" in the ultimatum game. The function of the word game is to motivate the assignment of roles; as Davis and Holt (1993 p 262) point out, it is important that the Proposer earns this right. The Proposer would suggest how an amount of \$10 should be split between the two of them (in multiples of \$0.50). The Responder got two alternatives, to accept in which case the proposal would be carried out, or to refuse whereupon none of them would receive any part of the 10 dollars.

Since the role of each subject was not yet known, there were natural reasons to ask each subject for his decision in both roles. At this second stage the subjects were asked about A/ the minimum he would require as Responder not to veto the proposal, and B/ how much he would keep as Proposer and how much he would propose the Responder as an ultimatum - take it or leave it. It is worth stressing that in this design, all the test persons made decisions in both roles instead of the more usual procedure of having each person to play the game in only one role.

Later the forms of each person were randomly paired with another and the specific roles set according to the scores in the word game. If the amount the Proposer wants to keep (P-value) and the amount the Responder demands as a minimum (R-value) stay within the 10 dollar limit, the ten dollar is divided according to the proposal of the Proposer (this implies that the Responder can get more than his R-value). If the \$10 limit is transgressed, none of the two person get any part of the \$10. The eventual gain of each person in the ultimatum game, was added to a base compensation of \$2.00 and delivered by mail.

### Table 1

The responses in the two roles were:

Proposer. (Average amount kept, dollars)

	Freshmen	Seniors	Total
Economists	6.30	6.02	6.15
Non-economists	5.65	5.20	5.44
			Total average 5.77 N = 92

Responder. (Average demanded minimum, dollars)

	Freshmen	Seniors	Total
Economists	1.38	1.98	1.70
Non-economists	2.85	1.98	2.44
			Total average 2.09 N = 92

Micro theory has a theoretical answer of special interest. According to that model the Responder should demand \$0.50. A claim for \$1.00 combined with an offer of \$0.50 dollar will result in getting nothing at all, so the rational demand is asking for just \$0.50 and hope for more. Since a rational Responder will ask for \$0.50, a rational Proposer will keep \$9.50 himself. A person behaving according to this theory will be labeled a "Micronic".

The expectation of the researchers were that economists should show a tendency to act according to micro theory, resulting in a higher dispersion between P and R numbers, while non-economists would show a tendency towards P and R equality. (A person choosing \$5 as Proposer and as Responder, will be labeled an "Egalitarian".) The numbers in the table support this trend. Carter and Irons conclude that the selection hypothesis was supported at the 2.5% level, while the learning hypothesis was discarded. The fact that the learning effect for Responder answers is significantly in the wrong direction is problematic.

There was also a stage in the Carter & Irons experiment that tested whether subjects could tell the solution according to micro theory (another part of this stage was two questions controlling that the subject had understood the basic rules of the game). Less than half of the students could give the correct answers, the economists being slightly better. Among freshmen economists a larger number suggested that a Responder should ask for \$5 than the correct \$0.50, and that a Proposer should keep \$5 instead of the correct \$9.50. People not only failed to behave according to micro theory, they even failed to understand the idea.

Carter and Irons conclude "Lastly, we note that while economists in our experiment behaved more in accordance with the rational/self interest model, this does not mean that their behavior was accurately predicted by the model."; "even economists sometimes fall short of the behavior expected of all good homines economici" (Carter & Irons 1991, p.177).

But even after acknowledging four respectively fifteen standard deviations from predictions, this is an understatement. Micro theory does not indicate a tendency but a specific solution. Since not one economist actually behaved according to micro theory, I hold that a lot more than the learning hypothesis falls. The result of the experiment undermines not only the claim of micro theory that most people should behave accordingly, but also the claim of the critics that micro theory say something (negatively) about the behavior of economists.

## **2 Two major problems with the critique of micro theory.**

I see two major problems with this analysis, and since these problems occur in a number of other reports (e.g. Frank et al. 1993), I think it is an issue of importance. The first is that the difference between the social groups is not explicitly connected to something of importance like a capacity of cooperation, but that such a bond is only vaguely assumed. The second problem is the lack of alternative theories - to be a supporter of "non-micro theory" is not a sufficient position. Let me elaborate a bit on those two questions.

When regarding the positions in the ultimatum game, the normative desirability of an equality tendency is in no way self-evident. The dispersion of the economists combines generosity as

Responder with selfishness as Proposer, while the orientation towards equality of non-economists combines generosity as Proposer with spite as Responder. The Responder with high demands is not only keen to improve his lot at the expense of somebody else, he is also ready to take a loss for himself to eliminate a larger benefit for somebody else. Most ethical philosophers from the classics to Rawls (1971) see selfishness as something legitimate if contained, but spite as unethical and not a legitimate base for social demands. From an ethical standpoint it is therefore hard to see an inclination towards equality (generosity as Proposer and spite as Responder) having more virtue than an inclination towards dispersion (selfishness as Proposer and generosity as Responder).

Neither from a practical point of view is there any good reason to assume a preference for an equal slant. Rather, it may be argued that a much stronger concept for civil cooperation has been dispersion built on performance, and that equality concepts such as "one for all and all for one" has mainly demonstrated its potential in military conflicts. There is no reason per se, to view the equality attitude as normal and desired, and a tendency versus dispersion as inferior.

To present an anomaly and raising doubts may be good enough given a strong theory with a close to 100 % explanatory power. However, the predictive weakness of micro theory has been documented many times before. Therefore one can see as a duty for the critics of micro theory to exert some constructive efforts, not only to gloat over its shortcomings. People negative to, and ignorant of economics, often harbor the idea that a company making no profit pursues a more socially desired activity than a profit-making one. The critics of micro theory come close to making similar crude assumptions regarding everything that associates with self-interest and compensation according to performance as negative. They therefore tend to, directly or indirectly, support anti-liberal and collectivist ideas. Benevolent manipulation is not an attractive choice itself, but if the alternative is perceived to be individualistic asociality, many moral claims for social obedience start to look quite beneficial (e.g. Dennet 1995 p 508 and Hofstadter 1985 p. 753). As is often the case, the descriptive and the normative are intimately connected, and with an unclear view of what is actually happening, the normative analysis becomes equally confused.

### **3 Further analysis - the introduction of the buffer**

In order to get more initiated understanding of the experimental results and whether more grounded normative comments can be made, I think one has to start with a broader analysis. As Sherlock Holmes remarked, it is of interest that the Baskerville dog did not bark. What choices did people have even if these were not frequently chosen? As mentioned Carter & Irons have provided me with their data on all the individual combinations and these say a lot more than average numbers. In the following I will present the whole range.

One possibility is to shift principles according to actual self-interest; thus to honor performance and ask for \$9.50 as Proposer, and to honor equality and demand \$5 as Responder. I label such a view "non-rational egoism" and define it as subjects whose sum of P- and R-values exceeds 10 dollars. Such selfish behavior was rare in this experiment, only

shown by two subjects out of 92. This result might be seen as supporting that a Rawlsian "veil of ignorance" can make people inclined to look for fair solutions.

A second category are those that demonstrate some "straight principle", providing a P and R sum of exactly ten dollars. There can be a number of different principles that fulfil this condition and both Egalitarians and Micronics belong to this category. My expectation was to find rather few of those two, but lots in the middle field, like a P-value of \$7 and a R-value of \$3. However, the striking result was that there were very few subjects in the whole category, just 8 people. (More specifically, there were 4 Egalitarians, and as mentioned none behaving according to micro theory.)

A third category are subjects with a sum in the range \$5 to \$9,50. Such a buffer between what they demand in the two roles and the \$10 can be seen as flexibility or conservatism, since it generates a good possibility of getting something. With 82 subjects, this constitutes the largest group and will be analyzed in more detail. But first, something about the fourth and last category.

Subjects with a sum of less than \$5 can be labeled "altruists". Keeping less than \$5 as Proposer is really to give away, because as the winner of the word contest they will hardly encounter claims of more than \$5. So keeping less is not just risk aversion - securing a return if unlucky to be paired against the demands of an Egalitarian. In the Responder position, it is hardly altruistic to raise claims, but to wait and see if the other person chooses to give. However, there were no subjects with combined numbers below \$5 in this test.

With no altruists, four Egalitarians, two non-rational egoists and no Micronics, it looks as if all major theoretical groups are endangered species. Let us get back to the majority in the buffer category and to see if some understanding can be obtained about what they are doing.

Micro theory states that the Proposer should expect that the Responder demands \$0.50. This is however not happening, and the question could then be revised to what is optimal to keep as Proposer under more realistic conditions. As can be seen in the table the average Responder asks for \$2.09, but there is a distribution around that number. For the Proposer, the marginal gain of getting more by increasing his keeping by one step is \$0.50. The marginal cost can be much higher, since he loses everything if trying to keep too much.

The Proposer has a choice between different alternatives  $P_i$  of P-value from \$10 to 0 in 21 steps (each of \$0.5) and the question is which alternative that gives him the highest expected value. Let  $F_i$  be the number of Responders accepting  $10 - P_i$  or less and  $N$  the total of number of participating Responders. A Proposer should try to maximize the expected value. The value is the product of the payoff  $P_i$  and the probability  $F_i / N$ . In the event of the probability  $1 - F/N$ , the payoff is 0 so the expected value only gets a contribution from the prior type of outcome. In order to illustrate the development of  $F$  and show if the probability is close to a linear function,  $f_i$  (defined as  $F_i - F_{i-0,5}$ ) is listed as the third column in the table below.

**Table 2**

$P_i$	$10 - P_i$	$f_i$	$F_i$	$P_i * F_i / N$	Exp. value
10	0	14	14	$10 * 14 / 92$	1.52
9.5	0.5	18	32	$9.5 * 32 / 92$	3.30
9	1	11	43	$9 * 43 / 92$	4.21
8.5	1.5	2	45	$8.5 * 45 / 92$	4.16
8	2	7	52	$8 * 52 / 92$	4.52
7.5	2.5	2	54	$7.5 * 54 / 92$	4.40
7	3	7	61	$7 * 61 / 92$	4.64
6.5	3.5	4	65	$6.5 * 65 / 92$	4.59
6	4	18	83	$6 * 83 / 92$	5.41
5.5	4.5	3	86	$5.5 * 86 / 92$	5.14
5	5	6	92	$5 * 92 / 92$	5.00

The result of this calculation is that the Proposer should keep \$6 and offer \$4 to the Responder since this alternative has the highest expected value. As can be seen in the table above the expected value for the Proposer amounts to \$5.41 when keeping \$6 and offering \$4, in contrast to \$3.30 if offering \$0.50. The keeping of \$6 is the rational choice of a Proposer given the actual distribution of Responder demands.

The average R-value in this experiment was \$2.09. If tying together the average R-value 2.09 and the optimal P-value 6.0 to the sum of \$10, the theoretical buffer becomes \$1.91.

If we now calculate the proposed amounts versus the average demand of the different groups, the following result emerges.

**Table 3**

Buffer between P-value of each group and average R-value of all participants

	Freshmen	Seniors	Total
Economists	1.61	1.89	1.76
Non-economists	2.26	2.71	2.47

Compared with the theoretical buffer of 1.91 the senior economists made the most rational offer. Two more groups lie within the +/- \$0.50 range and the fourth group is \$0.80 off. I think the fit is good enough to support a claim that the Proposers are securing maximum reward under real circumstances. The most critical comment towards economists would be that, even if the non-economists erred more severely, the freshmen economists erred on the "nasty" side that could result in less cooperation.

A problem when deciding what amount to keep is of course to estimate the demand curve. One way to get an idea is to judge others according to one self. This so called "attribution error" may be regarded as due to primitive information rather than as an error. In real life people are likely to interact with others like themselves, so therefore the buffer to average Responder demand of the own group is interesting.

**Table 4**  
Buffer between P-value and R-value of each group

	Freshmen	Seniors	Total
Economists	2.32	2.00	2.14
Non-economists	1.50	2.82	2.12

From these numbers the striking result is not the differences between the economists and non-economists, but the similarity. The learning effect seems to cause a decline in buffer for economists, but maturity causes an increase for non-economists. If nasty deviations are focused, the problem is with freshmen non-economists.

It is motivated to claim that the "statistical rationality" of the buffer model more deserves to be called "rational" than the "theoretical rationality" of game theory. The explanation can be seen as an empirical illustration of how to act according to decision theory as presented in classical works such as Raiffa (1968).

Camerer & Thaler (1995 p 214) comes to the same conclusion: "In ultimatum games Proposers act very much like sophisticated profit maximizers. They are sophisticated because they realize that unfair offers are likely to be rejected by the Responders. In the ultimatum game, it is the Responders' behavior that present the trickier (and more interesting) modeling problems."

#### **4 How to explain the Responder behavior?**

Responder behavior needs some explanation. How come they ask for more when this cannot improve their lot, that they simply risk some reward for themselves and more so for the co-player? Two things have to be taken into account, the reasons for spite and long term effects.

Economists and moral philosophers are in general agreement on spite. If my fellow man improves his lot, I might by egalitarian and/or selfish reason ask for a part, but I ought not to stop him making improvements just out of envy. If it is possible to improve the Pareto optimum, this should be done. Leaving these established views aside, seeing it instead from a practical point of view, a case for spite can readily be made. If a friend gets better off, that might be beneficial to myself, but improvements for enemies are evidently a disadvantage, and



improvements for people with whom I sometimes get in competitive situations, might also put me in a relative disadvantage.

The strongest case against spite is probably psychological. To hold everybody back in order for me not to stray behind is seldom possible and creates a destructive mood. However, without being a cynic it might be justified to note that even the capacity to share the fortunes of friends often seems to be quite limited. I think one can draw the conclusion from this experiment that spite was expressed by Responder numbers and expected by Proposer numbers. Worth noticing is that spite is not only shown by a small minority of bitter losers or as an effect of strict Egalitarian principles, but as a more general attitude. Sixty of the 92 subjects demanded more than \$0.50, that is, they were ready to make a sacrifice to stop the other person from a benefit (only 14 made no demands and asked for \$0).

From an evolutionary standpoint social behavior is preferably seen as profitable strategies. They are strategies that transform into rule of thumb, emotions indicating right and wrong, estimates of what others are up to etc. Most experiments try to create an artificial situation when an isolated action is asked for. No communication before, and anonymity afterwards are standard. This makes micro theory scientists expect short term rational decisions and non-micro theory scientists cling to unselfish reflexes. It looks like both groups are on the wrong tracks.

Experience has taught people that secrets become known and that effects on the personal reputation should be taken into consideration. One time anonymous Prisoner's Dilemma are of minor importance in real social life, but socially observed action and iterated Prisoner's Dilemmas are frequent. My explanation is that when occasionally taking part in an anonymous and isolated action, people behave partly as if in a more normal interaction. This means that they have to take into consideration effects upon the own reputation and its influence on future cooperation.

Many researchers want to develop experiments where these kind of external effects are completely eliminated. All kinds of deviation from the forecast of game theory are seen as results of some kind of "leaks". Students in the investigation might meet afterwards asking each other about what they did or how much they got, and test results might be very uniform so subjects can personally be linked to a certain behavior of a subgroup (an economist runs a risk of being seen as selfish, if economists generally score as selfish). Some foresight of such consequences might motivate people to adjust their behavior. If a completely separate and tight test environment could be created, people would at last behave according to game theory?

I doubt such a clinical separation is an interesting path for the development of tests. Much more interesting is the behavior that is not rigorously isolated, but is closer to the context of normal social behavior. Such a situation gives the actor some, but not total freedom from monitoring and policing. A mixed motive game has several advantages as a realistic setup for human social interaction compared to a zero-sum game or a cooperation game with no differences in interests. Likewise some communication and later information about the results of co-players adds parallelism to the game.

## 5 What is an optimal reputation?

In theory all straight principles give the same return, provided that individuals have equal capability and are as often in the Proposer as in the Responder position. But even when assuming equal capability there is a problem of reaching cooperation and that possibility might be influenced by offers and demands. What is then an optimal reputation, defined as giving best support for generating high value cooperation? Let me try to answer that by bringing up a critique to the radical policies of the Micronics and the Egalitarians. The former will have a lot of potential partners wanting him for the number 2 position, as Responder. According to the rule "something is better than nothing" he is contributing cheaply. When he is in a number 1 position, people will avoid him because he offers so little. In practice, the expected value for the Micronic will then have a tendency to get below average. In a similar way the Egalitarian will get dismal cooperation, but the other way around. When in number 1 position he will get many inter actors volunteering to make the minor contribution but getting an equal award, but when in the number 2 position the demands for equality are considered pretentious and potential inter actors will look for people with more reasonable demands.

For a person with above average capability there might still be some advantage being a Micronic. Even if he sometimes fails to get a partner for the number 2 slot, this might be compensated for by a larger proportion of potential number 1 situations. Such an attraction to a Micronic philosophy by the more capable is however an extra reason for the average inter actor of not bonding himself closely to a Micronic. The inter-actor will then often end up in the number 2 position with its dismal reward and seldom in the position highly rewarded.

The below average person will be more inclined to be an Egalitarian since it over compensates the more frequent number 2 position. Such a selection of principle according to selfish interest is an extra reason for the average inter actor to avoid also the Egalitarian; since he will do most of the work but only get half the reward.

Conclusively, Micronics and Egalitarians both have a reputation problem. Their generous offers will have more interested partners than they need, but their unattractive offers will be avoided if possible. Other inter actors will also be reluctant to take the unattractive and the attractive in a package by having a close relationship, since a Micronic will most often walk away with high proportion of reward and an Egalitarian will most often escape the main contribution.

## 6 Conclusion

A readiness to accept unattractive deals attract many cooperators, but not the cooperators you really want, those giving you a descent average. To show some pride might be good for screening them out, but at the same time it should not be exaggerated so that you will be regarded pretentious. Some low amount of spite might be the right message and it might be worth to abstain from some cooperation that judged in isolation would have been beneficial. In addition to rejecting sham suggestions like \$5-demands from Responders and \$0.50 offers

from Proposers, there are reasons to show flexibility. Even if the personal opinion of fairness is a P-value \$7 and R-value \$3, there is a point in not discarding a person who insists that fairness should be 8-2 or another person thinking a fair split should be 6-4. One strategy is to not only avoid being extremist, but also to exclude being "extreme middle". The subject emerging in this experiment is not the economic man of micro theory nor is he following other strict principles - he looks more like a trader ready to consider "all serious offers". But isn't that some kind of economic man?

My conclusion is that micro theory does not provide an explanation of a common way of thinking. People do not think in isolated actions, but expect observation and leaks - do they seem stubborn or weak in the observers' minds? Economists and non-economists seem to be doing the same thing in the game, they just have a minor difference in point of reference. I cannot see much of altruism and unselfish behavior. The game is better understood as played according to other rules and the subjects seem to have quite a good feel for what is going on. Before they were split into economists and non-economists they were all selfish group-living primates - and that is what they still are.

## Acknowledgement

I want to thank Ingolf Ståhl and Anders Westlund for valuable comments.

## References

- Camerer, Colin & Thaler, Richard. 1995. "Anomalies - Ultimatums, Dictators and Manners" in *Journal of Economic Perspectives* volume 9, number 2 pp 209-219
- Carter, John & Michael Irons. 1991. "Are Economists Different, and If So, Why?" in *Journal of Economic Perspectives*, 5 - 2, 171-177.
- Davis, D. & Holt, C. 1993. *Experimental Economics* Princeton University Press, Princeton NJ.
- Dennett, Daniel. 1995. *Darwin's Dangerous Idea*. Allen Lane. The Penguin Press, London).
- Frank, Robert H., Thomas Gilovich and Dennis T. Regan. 1993. "Does Studying Economics Inhibit Cooperation?" in *Journal of Economic Perspectives*, 7 - 2, 159-171.
- Hofstadter, Douglas. 1985. *Metamagical Themas: Questing for the Essence of Mind and Pattern*, Basic Books, New York.
- Marwell, Gerald and Ruth Ames. 1981. "Economists Free Ride, Does Anyone Else? Experiments on the Provision of Public Goods, IV", in *Journal of Public Economics*, vol 15 - 3, 295-310.
- Rawls, John. 1971. *A Theory of Justice*. Oxford University Press, Oxford.
- Raiffa, H. 1968. *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Addison-Wesley, Reading, Mass.