An Experimental Comparison of the Fairness Models by Bolton and Ockenfels and by Fehr and Schmidt*

Dirk Engelmann[†] and Martin Strobel[‡]
Humboldt-Universität zu Berlin

January 29, 2000

Abstract

In this note we present an experiment to compare the two competing fairness theories by Bolton and Ockenfels and by Fehr and Schmidt. For most experiments that these theories have previously been applied to, they make similar predictions so that it is difficult to compare their predictive accuracy. We designed a very simple experiment that induces opposing predictions made by both theories. If the maximization of total payoff is in line with the decision predicted by Fehr and Schmidt, almost all subjects decide accordingly, whereas if it is in line with the prediction by Bolton and Ockenfels, decisions are dispersed.

1 Introduction

Attention has recently been drawn to Bolton and Ockenfels' "Theory of Equity, Reciprocity and Competition" [forthcoming], and Fehr and Schmidt's "Theory of Fairness, Competition, and Cooperation" [1999], in the remainder of this note denoted by ERC and F&S, respectively. Both groups of authors quote an impressive number of experimental results that can be explained by their theories. However, if one wants to compare the predictive

*We thank Simon Gächter, Wieland Müller and Hans Normann for helpful comments. Financial support by the Deutsche Forschungsgemeinschaft through Sonderforschungsbereich 373 and by the Deutsche Girozentrale - Deutsche Kommunalbank is gratefully acknowledged.

†Institut für öffentliche Wirtschaft und Wirtschaftspolitik, Humboldt-Universität zu Berlin, Spandauer Str. 1, D-10178 Berlin, Germany, FAX: +49-30-2093 5787, phone: +49-30-2093 5786, e-mail: engelman@wiwi.hu-berlin.de

‡Institut für Wirtschaftstheorie III, Humboldt-Universität zu Berlin, Spandauer Str. 1, D-10178 Berlin, Germany, FAX: +49-30-2093 5704, phone: +49-30-2093 5724, e-mail: strobel@wiwi.hu-berlin.de

power of these two theories, most of the experiments that are interpreted retrospectively with these two theories are not helpful. For most experiments, they make equal or very similar predictions and thus their performance is similar. For example, both theories explain the results in ultimatum bargaining games reasonably well.

Both models rely on inequality aversion to explain deviations from selfish behavior. The fundamental difference between ERC and F&S is represented in the motivation or utility functions. The motivation function of ERC is given by $v_i(y_i, \sigma_i)$, with y_i denoting the own payoff and σ_i the share of the total payoff, and v_i for given y_i being maximal if $\sigma_i = \frac{1}{n}$, n being the number of players. F&S assumes a utility function $U_i(x) = x_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_j - x_i, 0\} - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_i - x_j, 0\}$ with $\alpha_i > \beta_i > 0$ and x_i the payoff of subject i.

Hence ERC assumes that subjects like the average payoff to be as close as possible to their own payoff while F&S assumes that subjects dislike a payoff difference to any other individual. Thus while according to ERC a subject would be equally happy if all subjects received the same payoff or if some were rich and some were poor as long as she received the average, according to F&S she would clearly prefer the first situation. In a real life situation F&S predicts that the middle class would tax the upper class to subsidize the poor, in an ERC world the middle class would just be satisfied.

To obtain explicitly opposite predictions by the two theories we chose a very simple game that focuses on their fundamental difference. Furthermore, it was completely reduced to the primary question that is considered by these theories, that of distribution. All a subject had to do was to choose an allocation of money between two other subjects, called persons in the experiment. The person who had the choice received an intermediate payoff and chose between three different allocations between a person who received in any case more than her and a person who in all allocations received less. These allocations were such that whenever she chose one with an average payoff for the other two persons closer to her's, both individual payoffs became more distant from her's, i.e. the richest person got richer and the poorest poorer. ERC predicts that she chooses the allocation that is most unequal between the other two persons because the average payoff is closest to her own. The opposite allocation is predicted by F&S to be chosen, since it minimizes the distance to the payoff of both the richer and the poorer person.

Apart from equity considerations other aspects might influence the decision of the subjects. In particular preferences for efficiency may lead to a choice that maximizes total payoff. (This would also result form altruism, but since subjects could maximize total payoff without any cost, in cannot be decided whether this is the actual motivation.) To prevent interference with the objective of our experiment, we designed two treatments, one in which following the prediction by ERC leads to a maximization of total payoff, one where maximization of total payoff is in line with the F&S prediction. The preferable way to prevent interference by efficiency would have been that all allocations yielded the same total payoff. However, if the own payoff is fixed, ERC only yields a clear prediction if the average and thus the total payoff of the other subjects differs in the three allocations.

We applied what could be called a mini strategy method. All subjects had to choose an allocation, while groups of three were randomly formed later on. Subjects knew that their decision would only matter if they were assigned to be person 2 in their group and that their payoff was fixed in that case, such that their decision could never influence their own payoff. Apart from generating three times the data by that method it also secured that all subjects were considered to be equally entitled to the money since they had all performed the same task. It also prevented that we had to pay subjects for doing nothing.

In the treatment where the prediction of F&S leads to a maximization of total payoff the results clearly confirm this prediction. The broad majority of subjects made that decision. In the other treatment subjects chose in about equal proportions the two extreme allocations and some chose the intermediate allocation. Hence in our simple decision task the performance of F&S is much better than that of ERC, although both theories ignore the importance that subjects assign to efficiency.

Section 2 presents our experimental design, followed by the experimental results in section 3. Section 4 concludes.

2 Experimental Design and Predictions

We conducted the experiment at the end of the lecture in one of the first weeks of an introductory economics course at Humboldt-Universität zu Berlin. 136 subjects took part in the experiment, 68 for each treatment. Each subject received a decision sheet with the instructions and a questionnaire, which we used to gather some biographical data and to check whether the subjects understood the task completely.

The decision sheet contained three different allocations of money between three persons, of which they had to choose one. They were informed that we would randomly form groups of three later on and would also assign the three roles randomly. Only the choice of that subject selected as person 2 mattered. The payoff of that subject was the same in all the three allocations they had to choose from. Hence no subject could influence her own payoff.

The allocations differed by the distance between the payoffs of person 2 and the other persons. The allocation where the distance between the payoffs of person 2 and each of the other persons was minimal was such that the distance to the average payoff was maximal and vice versa. Thus ERC and F&S predict the choice of exactly opposing allocations, since the former assumes that a disutility is caused by a difference between the own payoff and the average payoff, while the latter is based on a dislike for any payoff difference to other individuals.

Our two treatments differed by the effect the choice had on the total payoff. In treatment F the allocation that maximized the total payoff minimized the distance between the payoff of person 2 and those of persons 1 and 3, whereas it maximized the distance between the payoff of person 2 and the average of the payoffs of persons 1 and 3. Hence maximization of total payoff leads to the same decision as predicted by F&S (and that is why it is called treatment F). In treatment E there was the opposite relation, so that the choice predicted by ERC maximizes total payoff. Neither ERC, nor F&S, nor the maximization of total payoff predicts that the intermediate allocation will ever be chosen.

Each treatment was also divided into two subtreatments that only differed by the order in which the allocations were presented on the decision sheet. This was done to avoid some conceivable influence of a preference for the center or right allocation. (The allocation with intermediate payoffs was always presented on the left, since we considered this to be the most prominent position but it was the allocation we were not really interested in.) To avoid influence by computation errors we also noted the average payoffs of persons 1 and 3 and the total payoff for each allocation in the decision sheet, which looked similar to the tables below. (Actually, this implied that ERC was getting a pretty fair shot.) Subjects were paid in class the following week. They were identified by codes that were noted both on the decision sheets and on attached identification sheets that the subjects kept.

The possible allocations were for treatment F (subtreatment FH, in subtreatment FN columns B and C were reversed)

	A	B	C
Person 1	3.60 DM	4.60 DM	2.60 DM
Person 2	5.60 DM	5.60 DM	5.60 DM
Person 3	8.80 DM	8.20 DM	9.40 DM
Average 1 and 3	6.20 DM	6.40 DM	6.00 DM
Total 1,2,3	18.00 DM	18.40 DM	17.60 DM

and for treatment E (subtreatment EN, in subtreatment EH columns B and C were reversed)

	A	B	C
Person 1	3.20 DM	3.80 DM	2.60 DM
Person 2	6.40 DM	6.40 DM	6.40 DM
Person 3	8.40 DM	7.40 DM	9.40 DM
Average 1 and 3	5.80 DM	5.60 DM	6.00 DM
Total 1,2,3	18.00 DM	17.60 DM	18.40 DM

Treatment	ERC	F&S	inter
FN	2(6%)	29(85%)	3(9%)
FH	2(6%)	28(82%)	4(12%)
F	4(6%)	57(84%)	7(10%)
EN	14(41%)	12(35%)	8(24%)
EH	13(38%)	13(38%)	8(24%)
E	27(40%)	25(37%)	16(24%)
total	31(23%)	82(60%)	23(17%)

Table 1: Number of subjects choosing the allocations predicted by ERC (ERC), predicted by F&S (F&S), and the intermediate allocation (inter), by treatment and by subtreatment

In both these subtreatments F&S predicts a choice of allocation B, while ERC predicts a choice of allocation C. In subtreatment FH allocation B is efficient and in subtreatment EN allocation C.

3 Experimental Results

The results of our experiment are presented in Table 1. In both treatments there is virtually no difference between the two subtreatments. Hence we can conclude that the results are not driven by a preference for either the middle or the right column and we pool the data from the respective subtreatments. A preference for the left column might have increased the choices for the intermediate allocation, but since our interest is focussed on the two extreme allocations, we can ignore this possible though unlikely effect.

The results for treatment F are very clear. Of 68 subjects, 57 chose the allocation that led to a maximization of utility according to F&S and also to a maximization of total payoff. On the other hand, only 4 subjects chose the allocation predicted by ERC, and 7 subjects the intermediate allocation. A χ^2-test shows that the hypothesis that all three allocations are chosen with equal probability can be rejected ($\chi^2 = 78.21, p < .001$). Taking the number of choices of the intermediate allocation as given, a binomial test shows that the hypothesis that the allocations predicted by the two theories are chosen with equal probability can also be rejected ($p < .001$).

For treatment E the results are more dispersed. While of 68 subjects 27 chose the allocation predicted by ERC that also maximized total payoff, 25 decided according to the prediction by F&S and 16 subjects chose the intermediate allocation. The hypothesis that all three allocations are chosen with equal probability can not be rejected at a 5%-level according to a χ^2-test ($\chi^2 = 3.03, p > .2$). In particular we can conclude that in this treatment there is no significant difference between the probabilities with which the two extreme allocations are chosen.

Taking both treatments together, 82 subjects chose the allocation predicted by F&S, whereas 31 decided in line with ERC and 23 chose the intermediate allocation. According to a χ^2-test the probabilities with which the three allocations are chosen are significantly different ($\chi^2 = 45.19, p < .001$). A binomial test, taking the number of choices of the intermediate allocation as given, shows that the probabilities for the choices predicted by ERC and by F&S are significantly different ($p < .001$). Of the 136 choices in both treatments, 84 are in line with the maximization of total payoffs while 29 are opposite to it. A binomial test shows that this difference is significant ($p < .001$). Hence opposed to the assumption made by both ERC and F&S that efficiency does not matter we find a clear influence.

We find no notable gender effect on the probabilities of choices in line with ERC or F&S. However, the intermediate allocation is more attractive to males (16 choices) than to females (7 choices), with total numbers of males and females being almost equal (71 vs. 65).

If we exclude from the analysis those subjects who gave explanations for their decisions that indicated that they did not understand the instructions correctly, the results do not change substantially. For treatment E instead of 27 decisions in line with ERC and 25 with F&S, the numbers are 22 and 23, whereas those for the intermediate decision do not change. In treatment F instead of 57 decisions corresponding to the F&S prediction and 4 to ERC, we have 51 and 3, while the decisions for the intermediate allocation decrease from 7 to 6. The results of the statistical tests will not change if we use this adapted data.

In the explanation given for their decisions, 18 subjects explicitly referred to fairness. 17 of them chose according to F&S, including 8 subjects who also referred to the maximal total payoff. The remaining subject chose the intermediate allocation. Of 12 subjects who stated the reason for their decisions was maximization of total payoff (without explicit reference to fairness), 8 were in treatment F and thus chose the allocation predicted by F&S, the other four in treatment E chose according to ERC. Only one subject referred to the relative payoff but opposed to ERC aimed at maximizing the positive difference between the own and the average payoff. Thus among the subjects who explicitly mention fairness as a motivation F&S does much better than ERC.

The average payoff to each participant was DM 6.01. The experiment took about 20 minutes.

4 Conclusion

We presented a simple individual decision experiment to compare the relative performance of the competing fairness theories by Bolton and Ockenfels [forthcoming, ERC] and by Fehr and Schmidt [1999, F&S]. Subjects had to choose between three allocations for two other subjects with whom they were randomly chosen as a group. The allocations were chosen such that the two theories always predicted opposite decisions. The two treatments differed with respect to which theory was in line with efficiency.

The results are clearly in favor of F&S compared to ERC. In the treatment where the F&S prediction maximized total payoff, the vast majority chose this allocation. In contrast, in the other treatment choices split about equally between the F&S and the ERC prediction. In total the predictive power of F&S is significantly higher than that of ERC. Furthermore, efficiency plays a significant role.

Two main conclusions of our experiment arise. First, in such a simple decision task, subjects clearly care about equality of others' payoffs. Thus for this case, ERC is too simplifying since it ignores any influence of inequality between other persons. Second, subjects also clearly care for efficiency. Hence both F&S and ERC are too simplifying since they ignore in the utility or motivation function an efficiency (or possibly an altruism) component, that could explain the desire to maximize total payoff. As treatment E shows, the effects of inequality aversion concerning others' payoffs and of efficiency have roughly equal impact in the present experiment. The relatively high number of choices for the intermediate allocation and some of the explanations also indicate that subjects were driven by these two conflicting motives in this case.

Both the better accuracy of F&S and the relevance of efficiency may have been emphasized by the mini strategy method that we applied. Since subjects did not know the role they would end up in, they might have felt more sympathy with the other persons, than they might have felt if they had known to be person 2, although they knew their decision would only matter in that role. Increased sympathy with person 1 would yield more choices in line with F&S, while that with both other persons would imply choices maximizing total payoff.

A further possible limitation of both ERC and F&S is that they ignore intentions. However, since in the present experiment no choices of other players are observed, intentions do not matter and thus we cannot compare the performance of both theories to others that take intentions into account like Rabin [1993].

References

[1] Bolton, Gary E., and Axel Ockenfels, "ERC – A Theory of Equity, Reciprocity, and Competition", *American Economic Review,* forthcoming.

[2] Fehr, Ernst, and Klaus M. Schmidt, "A Theory of Fairness, Competition, and Cooperation", *Quarterly Journal of Economics,* CXIV (1999), 817–868.

[3] Rabin, Matthew, "Incorporating Fairness into Game Theory and Economics", *American Economic Review* LXXXIII, 1281–1302.