Working Paper

On the nature of fair behavior

Author(s):
Falk, Armin; Fischbacher, Urs; Fehr, Ernst

Publication Date:
1999

Permanent Link:
https://doi.org/10.3929/ethz-a-004372950 →

Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

forthcoming in: "Economic Inquiry"

Working Paper No. 17

On the Nature of Fair Behavior

Armin Falk, Ernst Fehr and Urs Fischbacher

August 1999

On the Nature of Fair Behavior*

Armin Falk, Ernst Fehr and Urs Fischbacher

University of Zurich**

Revised Version: August 1999

Abstract

This paper shows that identical offers in an ultimatum game generate systematically different rejection rates depending on the other offers that are available to the proposer. This result casts doubt on the consequentialist practice in economics to define the utility of an action solely in terms of the consequences of the action irrespective of the set of alternatives. It means, in particular, that negatively reciprocal behavior cannot be fully captured by equity models that are exclusively based on preferences over the distribution of material payoffs. Models that take into account players' fairness intentions *and* distributional preferences are consistent with our data while models that focus exclusively on intentions or on the distribution of material payoffs are not.

JEL-Classification: D63, C78, C91.

Keywords: Fairness, reciprocity, intentions, experiments, ultimatum game.

---

* Financial support by the Swiss National Science Foundation (Project 1214-05100.97) and by the MacArthur Foundation (Network on Economic Environments and the Evolution of Individual Preferences and Social Norms) is gratefully acknowledged. This paper is part of the EU-TMR Research Network ENDEAR (FMRX-CTP98-0238).

** Postal address: Institute for Empirical Research in Economics, Blümlisalpstrasse 10, CH-8006 Zurich. E-mail: falk@iew.unizh.ch, efehr@iew.unizh.ch, fiba@iew.unizh.ch. WWW: http://www.iew.unizh.ch/grp/fehr.

1. Introduction

There is by now considerable evidence that fairness considerations affect economic behavior in many important areas. In bilateral bargaining situations anonymously interacting agents frequently agree on rather egalitarian outcomes although the standard model with purely selfish preferences predicts rather unequal outcomes (Güth, Schmittberger and Schwarze 1982, Roth 1995, Camerer and Thaler 1995). In competitive experimental labor markets with incomplete contracts, fairness considerations give rise to efficiency wage effects that generate stable deviations from the perfectly competitive outcome (Fehr and Falk 1999). In several questionnaire studies (e.g. Bewley 1995, Campbell and Kamlani 1997) personal managers indicate that despite an excess supply of labor, firms are unwilling to cut wages because they fear that pay cuts are perceived as unfair and hostile by the workers and will, hence, destroy work morale. In principal-agent relationships reciprocally fair behavior causes a considerable increase in the set of enforceable contracts and, hence, large efficiency gains (Fehr, Gächter and Kirchsteiger 1997). To examine the forces that affect the perceptions of fairness and the determinants of fair behavior is, thus, not just of philosophical or academic interest.

A common feature of fair behavior in the above cited situations is, that in response to an act of party A that is favorable for party B, B is willing to take costly actions to return at least part of the favor (positive reciprocity), and in response to an act that is perceived as harmful by B, B is willing to take costly actions to reduce A's material payoff (negative reciprocity). This suggests that reciprocal behavior is an important component of fairness-driven behavior. Reciprocally fair behavior has been shown to prevail in one-shot situations and under rather high stake levels (Berg, Dickhaut and McCabe 1995, Roth, Prasnikar, Okuno-Fujiwara, and Zamir 1991, Cameron 1995).

In this paper, we show that identical offers in an ultimatum game trigger vastly different rejection rates depending on the other offers available to the proposer. In particular, a *given* offer with an unequal distribution of material payoffs is much more likely to be rejected if the proposer could have proposed a more equitable offer than if the proposer could have proposed only more unequal offers. Thus, it is not just the material payoff consequence of an offer that determines the acceptance but the set of available, yet

not chosen, offers is also decisive. This result not only casts serious doubt on the consequentialist practice in standard economic theory that defines the utility of an action solely in terms of the consequences of this action. It also shows that the recently developed models of fairness by Bolton and Ockenfels (forthcoming) and Fehr and Schmidt (forthcoming) are incomplete to the extent that they neglect "nonconsequentialist" reasons for reciprocally fair actions. These models assume that – in addition to their material self-interest - people also value the distributive consequences of outcomes. The impressive feature of these models is that they are capable of predicting correctly a wide variety of seemingly contradictory facts. For example: Why do competitive experimental markets with complete contracts typically converge to the predictions of the "selfish model" while in bilateral bargaining situations or in markets with incomplete contracts stable deviations in the direction of more equitable outcomes are the rule. However, despite their predictive success in important areas, our results indicate that there remain legitimate doubts whether these models capture the phenomenon of reciprocal fairness in a fully satisfactory way.

A parsimonious interpretation of our results, which is also suggested by psychological research, can be given in terms of "intentions".[1] Identical actions by the proposer are – depending on the available alternatives – likely to signal different information about the intentions of the proposer. Hence, if responders do not only take into account the distributive consequences of the proposers' actions but also the fairness of the proposers' intentions, their responses to identical offers may differ. Viewed from this perspective, our results suggest that fairness models should not only take into account that many people have preferences over the distribution of payoffs but also that many people value the fairness intentions behind actions. Models like this have been suggested by Rabin (1993), and Dufwenberg and Kirchsteiger (1998). However, as we will see, the recognition that intentions are important is not sufficient to account for our evidence because distributive concerns are important as well. Ultimately, it needs a model which combines both, preferences for distributive consequences and the role of intentions. An attempt in this direction is made by Falk and Fischbacher (1998).

[1] For a review of the psychological literature see Krebs (1970).

Before we present our experimental examination in detail, we would like to emphasize that the attribution of intentions for the evaluation of actions is not restricted to laboratory studies. We believe that it is also important in many real life situations. Take for instance the case that your neighbor caused a small damage at your car either intentionally or because of insufficient care. Most people would consider the intentionally caused damage as the more serious offence. Another important real life example that illustrates the importance of the attribution of intentions is the criminal law. It distinguishes carefully between criminal activities that are committed negligently and those committed with criminal intent. Similar distinctions are also made in the commercial law and the labor law. The punishment associated with a failure to meet obligations is, in general, dependent on judgments about the intention that caused the violation.

In the next section we describe our experimental design. Section 3 presents the results. The final section relates our findings to the literature and draws implications for theoretical modeling.

2. Experimental Design and Procedures

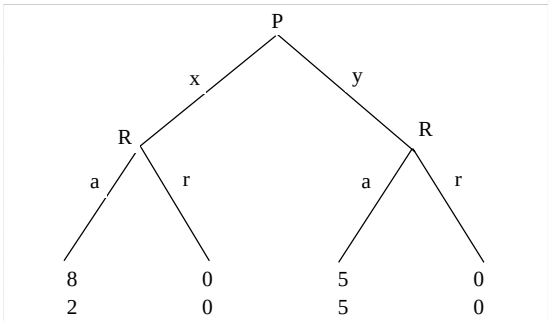
To examine whether identical offers trigger different rejection rates depending on the alternatives available to the proposer, we conducted four so-called mini-ultimatum games. Each of 90 experimental subjects participated in all four games. The mini-ultimatum games were extremely simple and share the same structure (see Figures 1a-d). In all games the proposer *P* is asked to divide 10 points between himself and the responder *R*, who can either accept or reject the offer. Accepting the offer leads to a payoff distribution according to the proposer's offer. A rejection implies zero payoffs for both players.

As Figures 1a-d indicate, *P* can choose between two allocations, *x* and *y*. In all four games the allocation *x* is the same while the allocation *y* (the "alternative" to *x*) differs from game to game. If *P* chooses *x* and *R* accepts this offer, *P* gets 8 points while *R* receives 2 points. In game (a) the alternative offer *y* is (5/5). This game is therefore called the **(5/5)-game**. Game (b) is called the **(2/8)-game** because the alternative offer *y* is to keep 2 points and to give 8 points to *R*. In game (c) *P* has in fact no alternative at all, i.e.,

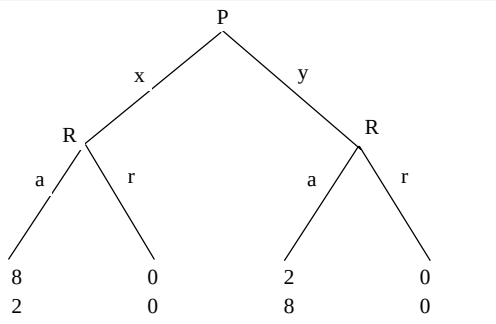
he is forced to propose an offer (8/2). We call it the **(8/2)-game**. Finally, in game (d) the alternative offer is (10/0), hence it is termed **(10/0)-game**. In order to get sufficient data we employed the strategy method. Every responder had to indicate his action at both decision nodes, i.e., for the case of an *x*- and for the case of a *y*-offer, without knowing what *P* had proposed.

At the beginning subjects were randomly assigned the *P*- or the *R*-role and they kept this role in all four games. Subjects faced the games in a random order and in each game they played against a different anonymous opponent. They were informed about the outcome of all four games, i.e., about the choice of their opponents, only *after* they had made their decision in all games. This procedure not only avoids income effects. It also rules out that subjects' behavior is influenced by previous decisions of their opponents. After the end of the fourth game subjects received a show-up fee of CHF 10.- plus their earnings from the experiment. For each point earned they received CHF -.80 so that in all four games together CHF 32.- (about $23) were at stake. The experiment lasted approximately 40 minutes. It was programmed and conducted with the software z-Tree (Fischbacher 1998).

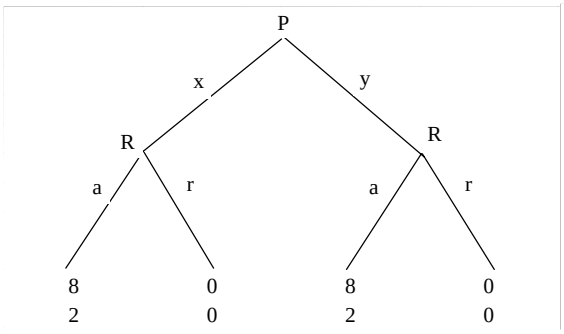
Figure 1: The mini ultimatum games



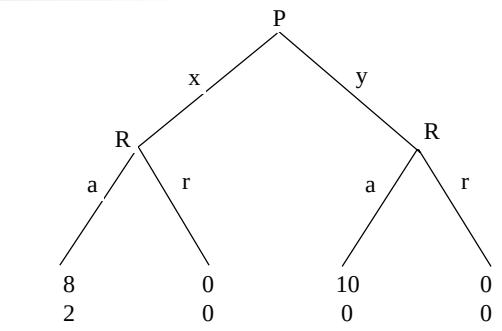
(a) (5/5)-game



(b) (2/8)-game



(c) (8/2)-game



(d) (10/0)-game

3. Predictions and Results

Since we are mainly interested in the variations of responders' behavior across the four games we shortly present the responder-predictions of the various models. The standard model with selfish preferences predicts that in all games the allocation (8/2) is never rejected. The Bolton-Ockenfels and the Fehr-Schmidt-model predict that the rejection rate of the (8/2)-offer is the same across *all* games. Since these models capture people's dislike for inequality, they are consistent with positive rejection rates. However, since they disregard that identical outcomes may be perceived as more or less fair, depending on the alternatives available to the first mover, they are not consistent with different rejection rates of the (8/2)-offer across the four games.

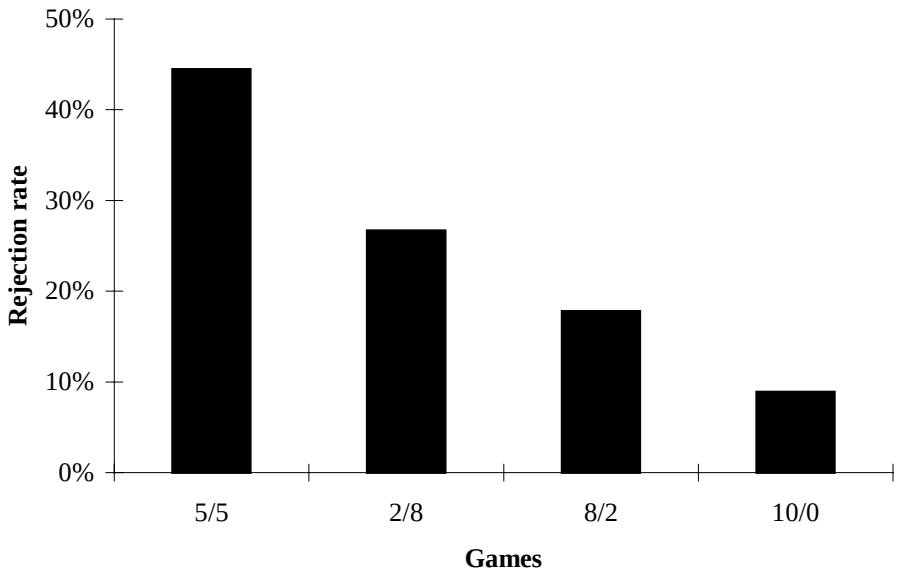
Intuitively, one would expect that in the (5/5)-game a proposal of (8/2) is clearly perceived as unfair because *P* could have proposed the egalitarian offer (5/5). In the (2/8)-game offering (8/2) may still be perceived as unfair but probably less so than in the (5/5)-game because the only alternative available to (8/2) gives *P* much less than *R*. In a certain sense, therefore, *P* has an excuse for not choosing (2/8) because one cannot unambiguously infer from his unwillingness to propose an unfair offer *to himself* that he wanted to be unfair to the responder. Thus, we would expect that the rejection rate of the (8/2)-offer in the (5/5)-game is higher than in the (2/8)-game. In the (8/2)-game *P* has no choice at all so that *P*'s *behavior* cannot be judged in terms of fairness. Responders can only judge the fairness of the *outcome* (8/2) and if they exhibit sufficient aversion against inequality they will reject this distribution of money. The rejection rate in the (8/2)-game measures, therefore, subjects' pure aversion against disadvantageous inequality. Since any attribution of unfairness to P's behavior is ruled out here we expect an even lower rejection rate compared to the (2/8)-game. Finally, offering (8/2) in the (10/0)-game may even be perceived as a fair (or less unfair) action so that the rejection rate of (8/2) is likely to be lowest in this game.

Figure 2 presents our main result. The bars represent the percentage of responders that reject the (8/2)-offer in the different games. The rejection rate in the (5/5)-game is highest (44.4 percent). 26.7 percent rejected the (8/2)-offer in the (2/8)-game, 18 percent in the (8/2)-game and 8.9 percent in the (10/0)-game. The non-parametric McNemar-test

confirms that the differences in rejection rates between the (5/5)-game and each of the other games is statistically significant ($p<.022$). The same test also confirms a significant difference between the (2/8)- and the (8/2)-game while the difference between the (8/2) and the (10/0)-game is not significant.[2]

These results indicate that pure aversion against inequality plays a role because 18 percent of responders reject the (8/2)-offer when *P* has no choice. This is evidence against pure intentions models like those of Rabin and Dufwenberg and Kirchsteiger. However, the results also clearly reject the implication of the Bolton-Ockenfels and the Fehr-Schmidt model that there are no differences in rejection rates across games. The rise in the rejection rate from 18 to roughly 45 percent in the (5/5)-game suggests that *intentions-driven* reciprocal behavior is a major factor.

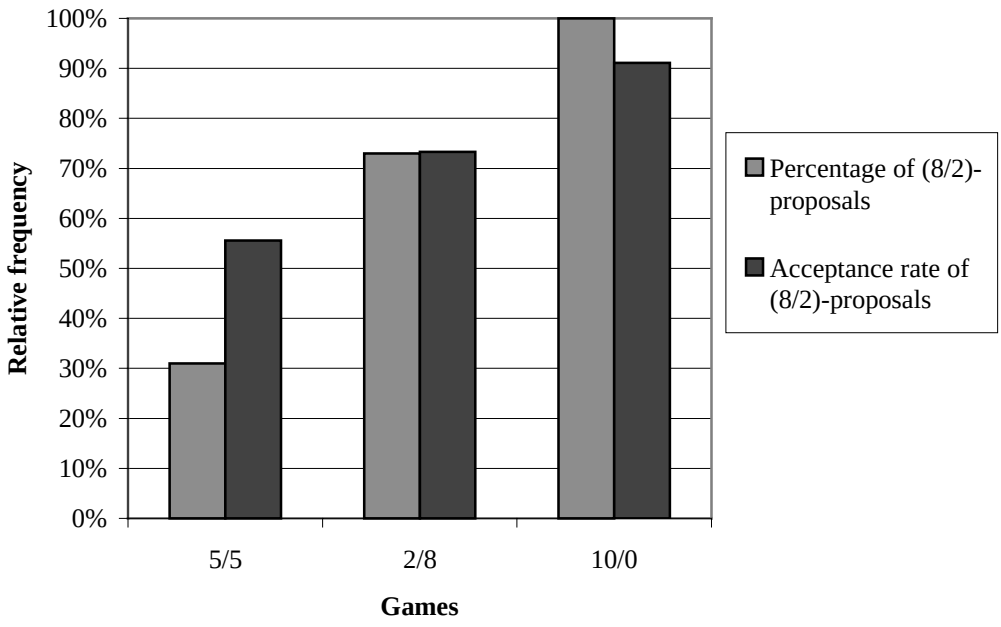
Figure 2
Rejection rate of the (8/2)-offer across games



[2] The rejection rates of the alternative offers (5/5), (2/8) and (10/0) are as follows. Nobody rejected the (5/5)-offer and only one subject rejected the (2/8)-offer. Almost 90 percent rejected the offer (10/0).

Finally, we take a look at proposers' behavior. Given the varying acceptance rates of the (8/2)-offers it is an interesting question whether proposers anticipated the differences in acceptance rates. Put differently, we can ask whether proposers realize that, given their alternative offers in the different games, responders have a higher tolerance to accept the (8/2)-offer. This would be further support for our interpretation that the fairness of an action is not only assessed by the distributive outcome but also by the intention that drives the action. Figure 3 provides the evidence. It shows that – when moving from the (5/5)-game to the (2/8)-game and then to the (10/0)-game – the monotonous increase in the acceptance rate of the (8/2)-offer is matched by a monotonous increase in the percentage of (8/2)-choices by the proposers.[3] The McNemar test indicates that the differences in the frequencies of the (8/2)-proposal are highly significant ($p < .001$).

Figure 3:
Percentage and acceptance of (8/2) proposals



[3] Since in the (8/2)-game proposers had no choice but to choose (8/2) such a comparison is meaningless for the (8/2)-game.

4. Concluding Remarks

The results of our experiment clearly show that the same action by the proposer in a mini-ultimatum game triggers very different responses depending on the alternative action available to the proposer. This suggests that responders do not only take into account the distributive consequences of the action by the proposer but also the intention that is signaled by the action. Supporting evidence for this interpretation is also provided by the experiments of Blount (1995)[4], Brandts and Sola (1998), and Güth, Huck and Müller (1998).[5] At a more general level our results also call into question the consequentialist practice in standard economic theory that defines the utility of an action solely in terms of the material consequences of this action.

The pattern of rejections in our experiment indicates that both distributional concerns and the attribution of fairness intentions are a driving force of reciprocally fair behavior and that models that focus exclusively on one aspect are incomplete. Therefore, the equity models of Bolton and Ockenfels and Fehr and Schmidt are not fully satisfactory because they have no explicit role for intentions while the pure intentions models of Rabin and Dufwenberg and Kirchsteiger are incomplete because they do not capture distributional concerns in a satisfactory way. Models that combine both driving forces of reciprocal behavior seem, therefore, most promising.

[4] The results of Blount (1995) may be affected by the fact that subjects had to make decisions as a proposer *and* as a responder before they knew their actual roles. After subjects had made their decisions in both roles, the role for which they received payments was determined randomly. In one of Blount's treatments deception was involved. Subjects believed that there were proposers although in fact the experimenters made the proposals. All subjects in this condition were "randomly" assigned to the responder role.

[5] There is also evidence for the view that distributive concerns alone as well as intentions alone cannot explain nice actions in response to nice behavior (see Charness 1996). For a dissenting view see, however, Bolton, Brandts and Ockenfels (1998).

References

Berg, Dickhaut and McCabe (1995): "Trust, Reciprocity, and Social History", Games and Economic Behavior 10, pp. 122-142.

Blount, S. (1995): "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences", Organizational Behavior and Human Decision Process 63, 131-144.

Bolton, G. and Ockenfels, A. (1999): A Theory of Equity, Reciprocity and Competition, forthcoming in American Economic Review.

Bolton, G. E., Brandts, J. and Ockenfels, A. (1998): "Measuring Motivations for the Reciprocal Responses Observed in a Simple Dilemma Game", Experimental Economics 1, 207-220.

Brandts, J. and Sola, C. (1998): "Reference Points and Negative Reciprocity in Simple Sequential Games", Discussion Paper, Universidad Autonoma de Barcelona.

Camerer, C. and Thaler, R. (1995): "Ultimatums, Dictators, and Manners", Journal of Economic Perspectives 9, 209-219.

Cameron L. (1995): "Raising the Stakes in the Ultimatum Game: Experimental Evidence From Indonesia", Princeton University, Industrial Relations Section, Working Paper: 345.

Charness, G. (1996): "Attribution and Reciprocity in a Simulated Labor Market: An Experimental Investigation", Discussion paper, University of Berkeley.

Dufwenberg, M. and Kirchsteiger, G. (1998): "A Theory of Sequential Reciprocity", mimeo, CentER for Economic Research, Tilburg.

Falk, A. and Fischbacher, U. (1998): "A Theory of Reciprocity", Discussion paper University of Zurich

Fehr, E. and Falk, A. (1999): "Wage Rigidities in a Competitive Incomplete Contract Market", Journal of Political Economy 107, 106-134.

Fehr, E., Gächter, S., and Kirchsteiger G. (1997): "Reciprocity as a Contract Enforcement Device: Experimental Evidence", Econometrica 65, 833-860.

Fehr, E. and Schmidt, K. (1997): "A Theory of Fairness, Competition and Cooperation", forthcoming in Quarterly Journal of Economics.

Fischbacher, U. (1998): "z-Tree. Zurich Toolbox for Readymade Economic Experiments." mimeo, University of Zurich.

Güth, W., Huck, S. and Müller, W. (1998): "The Relevance of Equal Splits – On a Behavioral Discontinuity in Ultimatum Games", Discussion paper, Humboldt University Berlin.

Güth, W., Schmittberger, R. and Schwarze, B. (1982): "An Experimental Analysis of Ultimatium Bargaining", Journal of Economic Behavior and Organization 3, 367-88.

Krebs, D. L. (1970): "Altruism – An Examination of the Concept and a Review of the Literature", Psychological Bulletin 73, 258-302.

Rabin, M. (1993): "Incorporating Fairness into Game Theory and Economics", American Economic Review 83, 1281-1302.

Roth, A., Prasnikar, V., Okuno-Fujiwara, M. and Zamir, S. (1991): "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study", American Economic Review 81, 1068-1095.

Roth, A. E. (1995): "Bargaining Experiments", J. E. Kagel and A. E. Roth (eds.): Handbook of Experimental Economics, Princeton University Press, Princeton.