ON THE NATURE OF FAIR BEHAVIOR

ARMIN FALK, ERNST FEHR, and URS FISCHBACHER*

This article 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. (JEL D63, C78, C91)

I. 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. 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 as shown in Fehr and Falk (1999). In several questionnaire studies, for example, in studies by Bewley (1999) and Campbell and Kamlini (1997), personnel 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. Fehr et al. (1997) show that in principal-agent relationships reciprocally fair behavior causes a considerable increase in the set of enforceable contracts and hence large efficiency gains. 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 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.2

In this article 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

1. See, for example, Güth et al. (1982), Roth (1995), or Camerer and Thaler (1995).

2. See Berg et al. (1995), Roth et al. (1991), or Cameron (1999).
the acceptance but the set of available, yet not chosen offers is also decisive. This result 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 (2000) and Fehr and Schmidt (1999) 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 correctly predicting a wide variety of seemingly contradictory facts. They predict, for example, why competitive experimental markets with complete contracts typically converge to the predictions of the selfish model, whereas 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 legitimate doubts remain as to 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. 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 take into account not only 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 take into account not only 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 that combines both preferences for distributive consequences and the role of intentions. An attempt in this direction is made by Falk and Fischbacher (1999).

Before we present our experimental examination in detail, we 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 small damage to your car either intentionally or because of insufficient care. Most people would consider the intentionally caused damage the more serious offense. Another important real-life example that illustrates the importance of the attribution of intentions is the criminal law. It distinguishes carefully between criminal activities committed negligently and those committed with criminal intent. Similar distinctions are also made in commercial law and labor law. The punishment associated with a failure to meet obligations is generally dependent on judgments about the intention that caused the violation.

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

II. 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 one of our 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 and allocation $y$ (the “alternative” to $x$) differs from game to game. If $P$ chooses $x$ and $R$ accepts this offer, $P$ gets 8 points and $R$ 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. Note that in the (2/8)-game P has only the choice between an offer that gives P much more than R (i.e., 8/2) and an offer that gives P much less than R (i.e., 2/8). In game (c) P has in fact no alternative at all, that is, he is forced to propose the offer (8/2). We call it the (8/2)-game. Finally, in game (d) the alternative offer is (10/0), hence it is termed the (10/0)-game.

To get sufficient data we employed the strategy method, that is, responders had to specify complete strategies in the game-theoretic sense. Thus, every responder had to indicate his action at both decision nodes, that is, 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 varying order, and in each game they played against a different anonymous opponent. They were informed about the outcome of all four games, that is, about the choice of their opponents, only after they had made their decision in all games. This procedure not only avoids income effects but also rules out that subjects' behavior is influenced by previous decisions of their opponents.

4. The payoff structure of this game is similar to the so-called best-shot game, which was first studied by Harrison and Hirshleifer (1989) and subsequently by Prasnikar and Roth (1992).

5. In principle, it is possible that the strategy method induces different responder behavior relative to a situation where responders have to decide whether to accept a given, known, offer. However, Brandts and Charness (2000) and Cason and Mui (1998) report evidence indicating that the strategy method does not induce different behaviors.
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 0.80, so that in all four games together CHF 32 (about US$23 at the time) were at stake. The experiment took approximately 40 minutes. It was programmed and conducted with the software zTree described in Fischbacher (1999).

III. PREDICTIONS AND RESULTS

Because we are mainly interested in the variations of responders’ behavior across the four games, we shortly present the responder-predictions of the various fairness 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 models predict that the rejection rate of the (8/2)-offer is the same across all games. Because these models capture people’s dislike for inequality, they are consistent with positive rejection rates. However, because 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.

The purely intention-based models by Rabin (1993) and Dufwenberg and Kirchsteiger (1998) are in principle compatible with different rejection rates for identical offers across games. The major reason for this is, however, that both models exhibit multiple equilibria. To be more precise, for each game Rabin’s model is compatible with the rejection and with the acceptance of the (8/2)-offer. Similarly, the Dufwenberg and Kirchsteiger model is compatible with the rejection and the acceptance of the (8/2)-offer in each of the first three games. We would like to stress, however, that a pure intention model, which formalizes the perceived unfairness of the intention as the only reason for rejecting an offer, should predict that no rejections occur if proposers cannot signal any intention. This is in our view the case in the (8/2)-game. In this game the proposer has no real choice and can therefore signal no intention. Thus, if only intentions matter, we should observe no rejections in the (8/2)-game.

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 subjects’ pure aversion against disadvantageous inequality. Because 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 the lowest in this game. The model by Falk and Fischbacher captures the essence of these intuitions. It predicts a positive rejection rate for the (8/2)-offers in all games and a higher rejection rate of the (8/2)-offer in the (5/5)-game, compared to the other games.

Figure 2 presents our main results. 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. Twenty of the 45 responders (44.4%) rejected the (8/2)-offer. Twelve subjects (26.7%) rejected the (8/2)-offer in the (2/8)-game, 18% in the (8/2)-game, and 8.9%
The rejection rate 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% rejected the offer (10/0).

7. 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% rejected the offer (10/0).

8. Because in the (8/2)-game proposers had no choice but to choose (8/2), such a comparison is meaningless for the (8/2)-game.

9. The Cochran Q-test indicates that the differences in the frequencies of the (8/2)-proposal across the three games are highly significant (p < .0001).
IV. 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 take into account not only the distributive consequences of the proposer’s action but also the intention signaled by the action. Supporting evidence for this interpretation is also provided by the experiments of Blount (1995). Brandts and Sola (2001), and Güth et al. (2001). The work by Offerman (2002) shows not only that the attribution of fairness intentions is an important determinant of punishment behavior in Ultimatum games but also that these attributions affect punishment behavior in other games as well.

At a more general level our results also imply that the utility of an action does not solely depend on the material consequences of the action but is also directly affected by the other available actions. This dependence has far-reaching consequences because it means that a decision maker can more easily enforce his or her preferred actions against opposition by secretly constraining the set of available actions or by pretending that certain actions are not available. At the theoretical level our results indicate that fairness models that are exclusively based on either distributional concerns or on the attribution of fairness intentions are incomplete. Therefore, the equity models of Bolton and Ockenfels as well as Fehr and Schmidt are not fully satisfactory because they have no explicit role for intentions, whereas 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, as those by Falk and Fischbacher (1999) or by Charness and Rabin (2002), are therefore most promising.

REFERENCES


