

KLAUS ABBINK, ABDOLKARIM SADRIEH and SHMUEL ZAMIR

## FAIRNESS, PUBLIC GOOD, AND EMOTIONAL ASPECTS OF PUNISHMENT BEHAVIOR

**ABSTRACT.** We report an experiment on two treatments of an ultimatum minigame. In one treatment, responders' reactions are hidden to proposers. We observe high rejection rates reflecting responders' intrinsic resistance to unfairness. In the second treatment, proposers are informed, allowing for dynamic effects over eight rounds of play. The higher rejection rates can be attributed to responders' provision of a public good: Punishment creates a group reputation for being "tough" and effectively "educate" proposers. Since rejection rates with informed proposers drop to the level of the treatment with non-informed proposers, the hypothesis of responder's enjoyment of overt punishment is not supported.

**KEY WORDS:** experimental economics, fairness, public goods, punishment, ultimatum bargaining

**JEL CLASSIFICATIONS:** C78, C91, C92, D82

### 1. INTRODUCTION

Typically, metropolitan taxi drivers and their passengers are involved in one-shot games. This is especially true in the interaction with non-local passengers, who will most probably not meet any cab driver more than once. Generally, the taxi driver not only has an informational advantage over the passenger, but also a bargaining advantage, since the passenger's cost of quickly finding a different means of transportation is often prohibitively high. Given these advantages and the one-shot nature of the game, we should expect to find taxi drivers practically dictating the terms of trade to passengers almost all of the time. This seems to be the reason why so many municipal authorities issue strict codes of conduct for taxi operators and attach severe sanctions to the breach of these rules. However, if even the slightest cost or effort is involved for the plaintiff in the



legal action against cheaters, we should not observe any such punishment action initiated by passengers in such one-shot games. Yet, casual observation suggests that cheated passengers occasionally do spend the cost and effort that is needed to report corrupt cab drivers to the authorities.<sup>1</sup>

The cheated taxi passenger is just one of many examples in which economic agents may actually choose a costly punishment action, without receiving an immediate monetary reward and without expecting to build a personal reputation that pays off in the long run. The choice of the action in such cases seems to be based on one of three motives: an intrinsic desire for justice, the enjoyment of overt punishment, or some form of strategic group reputation building. In terms of economic modeling, the former corresponds to the punisher's positive utility of punishing cheaters, regardless of the effect this punishment may or may not have on later payoffs. This can be referred to as the fairness aspect of punishment behavior. The second motive addresses the emotional satisfaction a punisher may have from knowing that cheaters know why and when they are brought to justice. In contrast to the first case, the punisher has a positive utility of punishing only if the punishment is overt and immediate (i.e., if the cheaters can unambiguously connect the punishment to this special case of their wrong-doing). The third motive corresponds to punishers caring about future payoffs of the group as a whole, perhaps also hoping that a more favorable "social environment" may enhance their own future payoffs, even if the direct effect on their own payoffs is negative. Obviously, this describes the public good aspect of punishment behavior, since "teaching" cheaters to "behave well" is a costly action that is taken to the benefit of the other group members. Testing for these three competing rationales of punishment behavior in naturally occurring one-shot games is practically impossible, because motives generally remain unobservable, even if data on choices and outcomes were accessible.

In this paper, we report an economic experiment that allows us to compare three of the competing motives of (costly) punishment action: the *resistance to unfairness*, the *enjoyment of overt punishment*, and the *contribution to group reputation*.

A no-feedback (*covered response*) treatment, in which overt punishment and group reputation building is impossible, is contrasted with a baseline full-feedback (*open response*) treatment. Since the experimental game is repeated in a number of consecutive rounds with new opponents, the dynamics of group reputation building can be identified in the behavior observed in the baseline treatment.

For the sake of simplicity, the interaction in our experiment takes part within the framework of the well-known *ultimatum game*. In this game, one player (the *proposer*, male pronouns) offers a second player (the *responder*, female pronouns) a division of a fixed cake. The responder can either accept or reject. If she accepts, the proposed division is implemented, if she rejects, both receive nothing. The subgame perfect equilibrium is straightforward: the responder does not reject any positive offers, since even the smallest positive payoff is better than receiving zero. In anticipation of this, the proposer offers no more than the smallest money unit and the responder accepts. Hence, under the common knowledge of rationality, the proposer virtually grabs the whole cake for himself.

Interestingly, and just as in the taxi example, observed behavior is different than theory predicts under the assumption of fully rational players, who are only motivated by their own monetary rewards. A large experimental literature, starting with Güth et al. (1982), has focused on this simple bargaining game. In almost all studies, dramatic deviations from the subgame perfect prediction are observed.<sup>2</sup> Typically, responders turn down small offers (a form of costly punishment) and proposers offer substantial amounts of money (very often up to an equal division of the cake).

Why do responders reject when this just means leaving money on the table? Such behavior is quite obviously not rational, if subjects are only concerned about own payoffs and the game is played only once. But, perhaps players do not only care about their own payoff, but also about the fairness of the outcome. In that case, they may reject low offers, because the rejection creates an allocation in which the proposer also receives nothing. In other words: Responders are willing to pay

the price of receiving nothing, because of their inherent *resistance to unfairness* (Thaler, 1988).

A number of models have recently been presented that incorporate inherent fairness motives of economic agents in different ways, these include the models by Bolton (1991), Rabin (1993), Dufwenberg and Kirchsteiger (1998), Fehr and Schmidt (1999), and Bolton and Ockenfels (2000). In these models, players are both concerned about their own payoff and the fairness of the outcome. Fairness is defined in various ways, but is always based on distributional issues and/or interaction history. The common denominator of these models is that they all apply standard methods of game theory to the interaction of agents with modified utility functions, i.e., incorporating some form of *fairness utility*. Generally, a remarkably good fit of model predictions to experimental data is reported by the authors. But, the players in these models are concerned about fairness mainly in a static manner. Though some authors acknowledge that the dynamic context matters,<sup>3</sup> the dynamic considerations of decision making in a repeated game context, which could motivate punishment behavior as an investment for future payoffs, are typically not captured in these models.<sup>4</sup>

The notion of punishment as an investment in future payoffs, however, gains some credibility when looking at another striking experimental observation: Although responders in ultimatum game experiments leave some money on the table by rejecting low offers, on average, their payoffs are much higher than in subgame perfect equilibrium play. This means that it is not necessarily only a fairness motive that is driving rejection behavior. Some kind of monetary motive may also be involved. In repeated play settings, such results are readily explained in terms of reputation building super-game strategies.<sup>5</sup> In the one-shot setting of the ultimatum game, however, reputation building is not possible, at least not on an individual level.<sup>6</sup> Both the proposer and the responder know that they will not meet and interact again.

While reputation on individual level cannot explain the success of responders in receiving relatively high average payoffs in the ultimatum game, perhaps reputation on group level

can. If the population of responders has a *group reputation* for being “tough” (i.e., for rejecting low offers), then proposers fearing rejections may increase their offers and, thus, drive responders’ payoffs up.<sup>7</sup> In this sense, the group reputation can be a public good for the population of responders and each individual responder can contribute to this public good by rejecting low offers.

There are good reasons to believe that responders will engage in the provision of the public good “group reputation.” First, there is abundant empirical evidence on the voluntary provision of public goods.<sup>8</sup> Second, a number of experiments by social psychologists show that the provision public goods is substantially higher when the benefits go to an *in-group*, where an in-group can be any group of people that share a group identity – even if this identity is induced arbitrarily by the experimenter.<sup>9</sup> In the case of the ultimatum game, any experimenter can confirm that the responders quickly develop a strong in-group feeling of being the unlucky underdogs. It is not hard to imagine that the strong group identification can lead to acts of solidarity, even if the individuals have no immediate contact to one-another.

The fundamental open question we address with this work is whether responders’ rejection behavior is driven by fairness motives, by the visibility of punishment to proposers, or by the desire to provide a public good, namely group reputation. Our experimental results strongly suggest that responders’ behavior is motivated both by a pure fairness motive, i.e., by the resistance to unfairness, and by the desire to create a more profitable interaction environment by contributing to their group’s reputation. The hypothesis that rejections may be driven by the enjoyment of the visibility of the punishment, however, finds no support in our data.

## 2. GAME

We use a reduced version of the ultimatum game, similar to what Bolton and Zwick (1995) introduced as the *cardinal*

*ultimatum game*, or what Gale et al. (1995) refer to as the *ultimatum minigame*. In our game, the proposer is restricted to two alternatives: the equal split (5, 5) or a division favoring himself (8, 2). The responder cannot reject the equal split. She can, however, reject the unequal offer, in which case both players receive nothing.

Compared to the standard ultimatum game, the reduced version used here allows simpler data analysis. All relevant data appear in one number for each role: subjects' behavior is expressed in rates of equal offers and responder rejection rates only. Previous experimental studies (Bolton and Zwick, 1995; Abbink et al., 2001) have shown that the reduced form ultimatum game captures the most relevant ultimatum game characteristics.

Each subject plays eight rounds of the game, all rounds in the same role (proposer or responder). Subjects are matched using a revolving (or round-robin) matching scheme, such that each proposer meets each responder only once and vice versa.

The game we use in the experiment is depicted (in extensive form) in Figure 1. It is the same for the *open* and the *covered* response treatment. In the open response condition, the proposer is informed of the responder's choice immediately after each round. In contrast, in the covered response condition the responder's choice is not reported to the proposer immediately. Only after the session is completed (i.e., after all eight rounds are played), proposers in the covered response treatment are

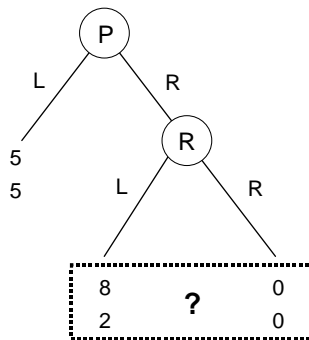


Figure 1.

informed about their earnings. From this information, they can derive how many times they have been punished, but not when and by whom.

The round-robin matching mentioned above is particularly important for the control condition with open responses. (To ensure compatibility, the same matching scheme is also used in the covered response treatment.) Due to the revolving matching scheme, a responder never meets the same proposer more than once, thus she cannot punish a proposer in the hope of getting a better offer from him in later rounds. Punishing a proposer, however, may have positive external effects on those other responders who meet the punished proposer in later rounds.

### 3. HYPOTHESES

This game with our two treatments allows us to test three plausible hypotheses on why people reject offers in the ultimatum game. The first hypothesis we test is that of *resistance to unfairness*, which captures the motivational assumptions underlying fairness utility models. These models are insensitive to the dynamic considerations that may play an important role in our experiment. Only the fairness of the final outcome – next to the own payoff – enters players' utility or evaluation functions. Since the responders have exactly the same choices and influence the final allocation in exactly the same way in both treatments, the observed distributions of final allocations should be indistinguishable in both treatments. Especially, the fact that the proposers in the covered response treatment are not informed on who punished them when should not make a difference for the choices the responders make.

**Hypothesis 1** (resistance to unfairness). *Rejections are motivated by the fairness of the final outcomes. No difference in rejection rates between the covered and the open response treatment are expected.*

The second hypothesis addresses a *group reputation* motive. Responders contribute to their group's reputation for being

“tough” by punishing proposers who make unfair offers. Since a proposer faces a much greater damage after the rejection of a low offer than the rejecting responder does, proposers are likely to move towards more generous offers in the face of a “tough” group of responders. Thus, the rejection of low offers can *educate* the proposer. In a “meet only once” context, the rejection is not meant to enhance the rejecting responder’s payoff immediately by educating the particular proposer for later occasions when the two meet again. Instead, educating proposers is a system of mutual contributions to a public good, namely to the group reputation.<sup>10</sup>

Obviously, group reputation can only be built up in the open treatment. This means that more rejections should be observed in the open treatment than in the covered treatment. But, as the experiment proceeds, the positive external effect of punishment decreases, because fewer and fewer encounters are left in which responders can profit from previous educative rejections by their fellow group members. Hence, if group reputation is the predominant motive of punishment behavior, then we should expect rejection rates to decline in the open treatment towards the end of the experiment. In the last round, there should be no difference between the open and the covered treatment.<sup>11</sup>

**Hypothesis 2** (contribution to group reputation). *Rejections are motivated by group reputation considerations. No rejections are expected in the covered response treatment. In the open response treatment, high rejection rates are expected in early rounds. These rates are expected to decline towards the end of the experiment, as the advantage of the group reputation diminishes.*

The third hypothesis addresses an *emotional* aspect of rejection behavior.<sup>12</sup> On the one hand, responders who are confronted with unfair offers might enjoy the overt and immediate punishment that is possible in the open treatment. On the other hand, when punishment is veiled and delayed as in the covered treatment, rejecting unfair offers is conjectured to cause no such enjoyment. This means that if the enjoyment of overt punishment is effective, then rejection rates must be



higher in the open than in the covered response treatment. However, the proposed effect is static, so that rejection rates should not be expected to decline towards the end of the experiment. Note, that this difference in the predicted development of the rejection rates distinguishes hypotheses 2 and 3.

**Hypothesis 3** (enjoyment of overt punishment). *Rejections are motivated by enjoyment of immediate punishment. No rejections are expected in the covered response treatment, because proposers making unfair offers cannot be punished overtly and immediately. In the open response treatment, high rejection rates are expected throughout the experiment (constant across all rounds).*

Table I summarizes the hypothesis and their predictions in the treatments.

#### 4. EXPERIMENTAL PROCEDURE

The experiments were computerized, with software developed using *RatImage* (Abbink and Sadrieh, 1995). The game was presented to the subjects in the game tree form, where the decisions were submitted by buttons at the branches of the tree. After a decision, the chosen branch was highlighted, the other

TABLE I  
The three hypothesis of our experiment.

	Hypothesis 1 (Fairness)	Hypothesis 2 (Reputation)	Hypothesis 3 (Enjoyment)
Covered response treatment	Rejection rates $> 0$	No rejection	No rejection
Open response treatment	Rejection rates $> 0$	Rejection rates $> 0$	Rejection rates $> 0$
Predicted pattern of rejection rates	No difference in rejection rates between treatments	Rejection rates decline over round in open response	Rejection rates stay constant over rounds in open response

ones lowlighted. In the open response treatment, the responder's choice and the resulting payoffs were immediately marked on the proposer's screen. In the covered response treatment, the proposer's screen only displayed a question mark between the two possible outcomes of that branch.

Six sessions of each treatment were conducted, with a total of 192 subjects. The experiments were run in the *RatioLab* at the Hebrew University of Jerusalem, Israel, and in the *Laboratorium für experimentelle Wirtschaftsforschung* at the University of Bonn, Germany. In a two-country experiment, the fact that the subjects in different countries are drawn from distinct subject pools must be taken into account. Differences in behavior might arise from country-specific cultural environments, but also from a different composition of the subject pools with respect to educational background, gender, age, majors of study, and others. In fact, in the four country study by Roth et al. (1991), the Israeli subjects' behavior was slightly different from behavior in the other countries. In Jerusalem, lower offers and a higher tendency to accept low offers were observed in the ultimatum game. Thus, the possibility of subject pool differences must be taken into account. We controlled for this possibility by splitting the two treatments between the two locations evenly. Three sessions of each treatment were run in each laboratory. Hence, if subject pool differences should exist, they cannot be in conflict with treatment differences.

All subjects were volunteers and were only given monetary incentives. The sessions started with an oral presentation of the written instructions (about 10 min) that were handed out to the subjects (reproduced in Appendix A). The sessions were conducted in the local language of each country. Much care was taken to ensure that the instructions were equivalent in both countries. The instructions were first written in English and then translated into each local language by a member of the local laboratory team. In the next step, another team member translated the translation back into English, without seeing the original English text. This procedure was repeated until the back-translation and the original text had converged. Convergence, however, was reached rather quickly, since a very neutral

and technical phrasing was used. The same method was applied to the screen output, which was also displayed in the local language of each country. The main decision screens are reproduced in Appendix A.

The social composition of the two subject pools was similar, since both laboratories are located in social science buildings and both subject pools mainly consisted of students of social sciences. The exchange rates of points to cash were adjusted in the two countries in a way that total earnings were comparable, in terms of teaching assistants' average hourly wage rates at each location. The exchange rate of DM 0.50 (roughly US\$ 0.33 at that time) per point was used in Bonn, and NIS 0.75 (about US\$ 0.25 at that time) was used in Jerusalem. Since the experiments were short (about 30–45 min in total), the resulting payoffs were on average well above the typical student's per hour wage.

## 5. RESULTS

First, we focus on the responder behavior. We begin with the analysis of aggregate rejection rates over the sessions. The raw data of the experiment are reproduced in Appendix B.

### 5.1. Responder rejection rates

Table II shows the average overall rejection rates (rejected unequal offers to total unequal offers) in the six sessions of each treatment, ordered from the lowest to the highest. The Jeru-

TABLE II  
Average rates of rejected unequal offers in each session (%).

Covered	6.5 <sup>+</sup>	17.6*	22.7 <sup>+</sup>	26.1*	31.0*	31.3 <sup>+</sup>	Avg. over all covered = 22.5
Open	10.0 <sup>+</sup>	20.0*	40.0*	43.5*	50.0 <sup>+</sup>	80.0 <sup>+</sup>	Avg. over all open = 40.6

saalem sessions are marked with a plus sign (+), the Bonn sessions with an asterisk (\*).

The rejection rates show that, even when the response is not reported to the proposer (covered treatment), almost one quarter of all unequal offers are turned down. This supports hypothesis 1, since responders exhibit resistance to unfairness that is entirely independent of all considerations of own monetary payoff maximization. Responders cannot openly punish proposers in the covered response treatment – not even as a group – in order to receive higher payoffs later. Thus, the relatively high average rejection rate (22.5%) is evidently motivated by negative emotions towards unfair actions or distributions. The disutility of unfair outcomes that in some way or another is contained in all fairness utility models (either by payoff comparison or otherwise) can be interpreted as a formalization of such a motive.

**Observation 1.** In the covered response treatment, substantially positive rejection rates are observed. This supports hypothesis 1 (resistance to unfairness).

However, resistance to unfairness cannot explain a different aspect of our data. There is a significant difference in the rejection rates of the two treatments. The average rejection rate in the open response sessions is 40.6%, which is about 75% higher than those of the covered treatment. The difference is significant with a  $p$ -value of 0.090 (one-tailed), according to the Mann-Whitney U-test applied to the average rates of rejection in sessions.<sup>13</sup> The pronounced difference between the rejection rates of the two treatments is consistent both with Hypothesis 2 and with Hypothesis 3.

**Observation 2.** In the open treatment, unequal offers are rejected significantly more often than in the covered treatment. This is consistent with Hypothesis 2 (contribution to group reputation) and with Hypothesis 3 (enjoyment of overt punishment).

Observation 2 reveals that next to the resistance to unfairness some other motive for rejecting unfair offers must be present in the open treatment. Comparing the aggregate rejection rates only does not allow a differentiation between the group reputation and the overt punishment motives of Hypotheses 2 and 3. In order to make such a distinction, we must take a closer look at the development of the rejection rates. While the enjoyment of overt punishment motive (Hypothesis 3) predicts rejection rates in the open treatment not to show a trend over experimental time, it is consistent with the group reputation motive (Hypothesis 2) that these rates decline, eventually reaching the level of rejections in the covered treatment.

Figure 2 shows the aggregate rejection rates over the eight rounds of the experiment. The differences between the rejection rates across treatments diminish towards the end of the session: in the last two rounds, when contributing to group reputation makes little or no sense, almost no difference between the treatments can be observed. Statistical tests of the last round rejection rates do not detect any significant treatment differences. This supports Hypothesis 2, but is not in line with Hypothesis 3.

The fact that average rejection rates get closer to one another is mainly due to the tendency of the rejection rates in the open response treatment to fall over time. From the first to the

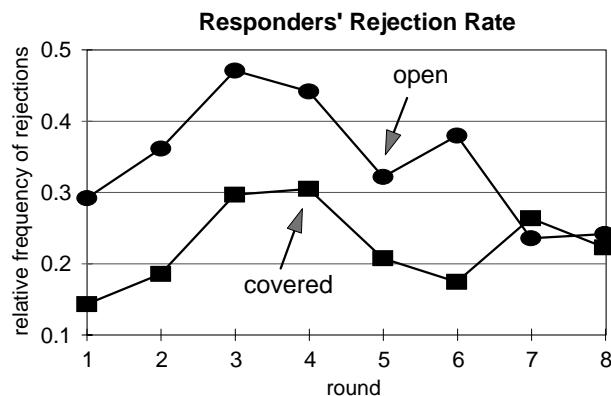


Figure 2.

second half of a session (four rounds each), the aggregate rejection rate falls from 39.8 to 29.1% in the open response treatment. The decrease is significant at  $p = 0.015$ , according to the Wilcoxon matched pairs signed rank test, applied to the difference of rejection rates between the first and the second half of the experiment in the six independent sessions. In contrast, the rejection rates in the covered response treatment fall only slightly (from 24.2 to 21.4%), and the decrease is not significant. The corresponding numbers for the single sessions are shown in Table III.<sup>14</sup>

**Observation 3.** In the covered treatment, no trend can be detected in responder rejection rates. In the open treatment, responder rejection rates fall significantly from the first to the second half of the experiment, reaching the level of rejections in the covered treatment in the last round. This supports Hypothesis 2 (contribution to group reputation), but is not in line with Hypothesis 3 (enjoyment of overt punishment).

The three observations above lead to the following conclusion: The fairly stable rate of rejections in all rounds of the covered

TABLE III  
Rejection rates in the first and the second half of the experiment.

Session	Open treatment			Covered treatment		
	Rounds 1–4	Rounds 5–8	$\Delta$	Rounds 1–4	Rounds 5–8	$\Delta$
1*	0.44	0.43	-0.01	0.38	0.23	-0.15
2*	0.18	0.21	+0.03	0.33	0.29	-0.04
3*	0.44	0.40	-0.04	0.19	0.17	-0.02
4 <sup>+</sup>	0.52	0.43	-0.09	0.20	0.25	+0.05
5 <sup>+</sup>	0.82	0.75	-0.07	0.29	0.33	+0.04
6 <sup>+</sup>	0.18	0.07	-0.11	0.13	0.00	-0.13
Over all	0.40	0.29	-0.11	0.25	0.21	-0.04

\*Bonn session.

<sup>+</sup>Jerusalem session.

response treatment reflects a basic rate of rejections due to resistance to unfairness in our subject pool. The higher rates of rejection in the early rounds of the open response treatment appear to incorporate some amount of rejections motivated by group reputation in addition to the basic rate of rejections due to resistance to unfairness.

### 5.2. Rates of equal offers

In Figure 3, the round by round aggregate rates of equal offers are depicted. Table IV shows the session averages of the rates of equal offers, ordered from the smallest to the largest. Again, Jerusalem sessions are marked with a plus sign (+), and Bonn sessions with an asterisk (\*).

In the covered response treatment, we observe equal offers in half of the cases (49.8%), which is about 40% higher than the 35.7% in the open response treatment. Aggregated over all eight rounds, the difference across treatments is significant at  $p = 0.026$  (one-tailed), according to the Mann–Whitney U-test, applied to average rates of equal offers in the sessions.

Already in the very first round, we observe a higher rate of equal offers in the covered response treatment. 70.8% of the proposers choose an equal offer in the first round, whereas

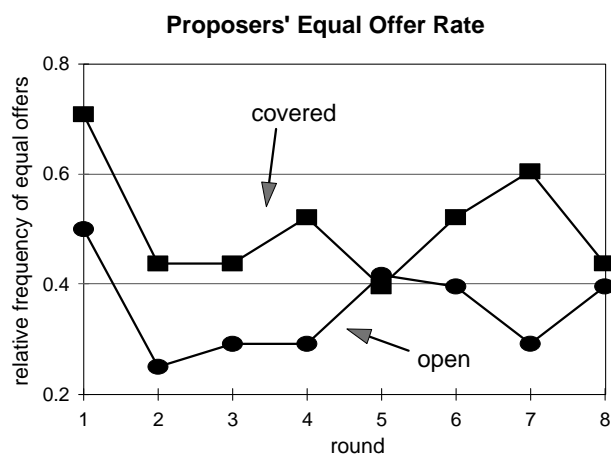


Figure 3.

TABLE IV  
Average rate of equal offers in each session (%).

Covered	31.3 <sup>+</sup>	46.9*	50.0 <sup>+</sup>	51.6 <sup>+</sup>	54.7*	64.1*	Avg. over all covered = 49.8
Open	21.9*	21.9*	21.9 <sup>+</sup>	28.1*	43.8 <sup>+</sup>	76.6 <sup>+</sup>	Avg. over all open = 35.7

50.0% do so in the open response treatment. According to Fisher's exact test, this difference is significant at  $p = 0.019$  (one-tailed). Note that the individuals' equal offers are independent in the very first round of the experiment.

**Observation 4.** The rates of equal offers are significantly higher in the covered response treatment than in the open response treatment.

Note that the rates of equal offers in the covered response treatment are significantly higher than in the open treatment, although the rejection rates of unequal offers are lower. This may seem a puzzle if we assume that proposers can foresee the lower rejection rates in the covered treatment perfectly. However, if we – more realistically – assume that proposers are not perfectly informed on the rejection rates they face, then an “exploration” hypothesis can explain the observed pattern of behavior. Subjects proposing in the open response treatment tend to test responders' propensity to reject by selecting the unequal offer branch. Learning from success and failure, however, is completely impossible for proposers in the covered response treatment. Thus, the lack of feedback seems to decrease the tendency to explore the responders' propensity to reject and to lead to the significantly higher rates of equal offers in the covered response sessions. In this sense, the proposers in the covered treatment tend to take on smaller risks than those in the open treatment.

It seems plausible that the high rates of equal offers in the covered response treatment are therefore due to an enhanced risk avoidance of the proposers. If a proposer makes an



unequal offer, his payoff can vary between 0 and 8, depending on the responder's choice. In contrast, he is guaranteed a payoff of 5 when making an equal offer. Thus, if the proposer believes that there is a positive probability for some responders to reject the unequal offer, the payoff on the right branch is uncertain, with an unknown probability, whereas the payoff on the left branch is certain.

### 5.3. Is group reputation effective?

Given that responders in the open response treatment use punishment to building up a group reputation for being "tough," the question arises whether their attempt is successful. Do early rejections actually induce proposers to shy away from the unequal offer, because of the expectation of punishment? If the answer is "yes," then we should observe increased rates of equal offers in those proposer populations that faced the highest rejection rates. For each session, we examine the rejection rates in the first half of the experiment and the change in the rate of equal offer from the first to the second half. We compute the correlation of these two measures over the sessions. If group reputation building is effective, then a tendency for high first half rejection rates to be followed by high increases in the rate of equal offers should be observed.

Figure 4 shows that, in fact, there is a strong correlation in the predicted direction. The greater the responders' reluctance to accept unfair offers, the greater is the rise in the frequency of equal offers with time. The Spearman rank correlation coefficient of  $r_s = 0.93$  is significantly different from zero at  $\alpha = 0.05$  (one-tailed). Thus, high frequencies of rejections correlate to higher increases in the rate of equal offers.

We now examine whether the positive correlation between rejection rates and the change in the rates of equal offers can be detected on an individual level. To see this, we check each round in which a proposer switches from a preceding round's unequal to a current round's equal offer. We count how often the equal offer was made after the last round's unequal offer was rejected, and compare this to the frequency of equal offers

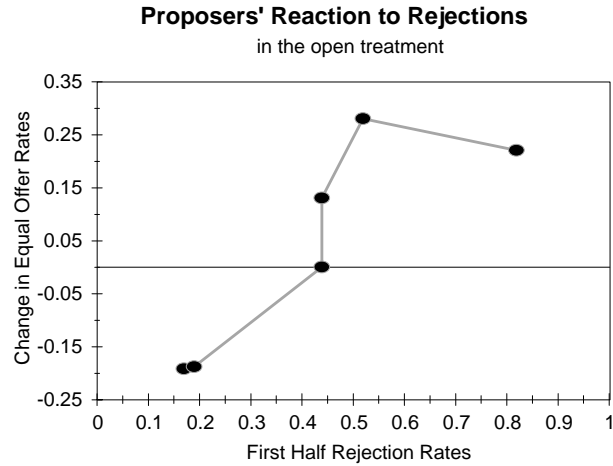


Figure 4.

following an accepted unequal offer. If proposers were successfully educated with rejections, then we should observe systematically more switches after punishments than after accepted unequal offers.

We find that the group reputation building is actually effective on the individual level. Observing a rejection, i.e., being punished, influences the proposers' behavior. Proposers' propensities to choose the equal offer is rather low if their unequal offer in the last round was accepted (16 out of 140 cases, relative frequency 11.4%). But, if punished, their tendencies to switch is almost four times higher (35/79, relative frequency 44.3%). This is consistent with the hypothesis that proposers in the open response treatment attempt to test the probability of rejection and switch to the equal offer if that probability is perceived as high.<sup>15</sup>

The observed pattern of switches points in the same direction in all six sessions of the open response treatment. Table V shows the frequencies of switches after punishments and after accepted unequal offers. The last column shows the difference between the relative frequencies of switches. In all six sessions this measure has a positive value. Thus, a switch occurs relatively more often after rejected than after accepted unequal offers. The matched-pair sign test rejects the null hypothesis

TABLE V  
Frequency of switches from the unequal of the equal offer.

Session	Switch after rejection (1)	Rejections rounds 1–7 (2)	Switch after accepted offer (3)	Accepted offers rounds 1–7 (4)	Difference in rel. frequency (1)/(2)–(3)/(4)
1	7	17	3	24	+0.29
2	2	10	2	33	+0.14
3	5	18	3	26	+0.16
4	8	17	4	17	+0.24
5	11	12	1	3	+0.58
6	2	5	3	37	+0.32
$\Sigma$	35	79	16	140	+0.33

that a switch is equally likely after punishment as after an accepted unequal offers with at  $p = 0.016$  (one-tailed).

**Observation 5.** In the open treatment, proposers tend to switch to the equal offer significantly more often after having observed a rejection than after accepted unequal offers.

After having switched to an equal offer, proposers often switch back. Of all equal offers, 46% are followed by an unequal offer. Thus, although proposers do react to punishment, they also show a strong tendency to switch back to the unequal offer. This implies that high rejection rates are necessary for a sustained group reputation effect on the responders side. Interestingly, the frequency of switching from the equal to the unequal offer is almost the same in both treatments: under the covered condition we observe a relative frequency of 47%.

#### 5.4. Are contributions to group reputation profitable?

We have found evidence that contributions to the group reputation are effective in the sense that proposers are influenced towards making more equal offers, after having observed rejections. However, since rejections are costly for the

responders, the effectiveness of such educative punishment does not immediately imply that such behavior is profitable for responder populations. In this section, we analyze how rejection rates and responder profits are distributed over the populations.

Figure 5 shows the relationship between the rejection rates in a session and the average responder payoff (in per cent of the cake) in that session. The numbers are computed on the basis of all eight rounds of the session. If contributing to the group reputation is profitable, then we should observe higher responder payoffs in the sessions with high rejection rates. The correlation between responder payoff and rejection rates seems to be U-shaped rather than monotonously rising or falling. It seems that the more extreme patterns of behavior are more profitable than those in between. Contributing to the group reputation seems to be profitable only if the rejection rates are sufficiently high to induce a strong effect. Where the rejection rates are intermediate, the responders pay the price for punishment, but the frequency of unequal offers is not sufficiently decreased to compensate the costs of punishment. To illustrate this point, we have included the vertical line into the figure. At a rejection rate of 0.375, expected payoff maximizing proposers

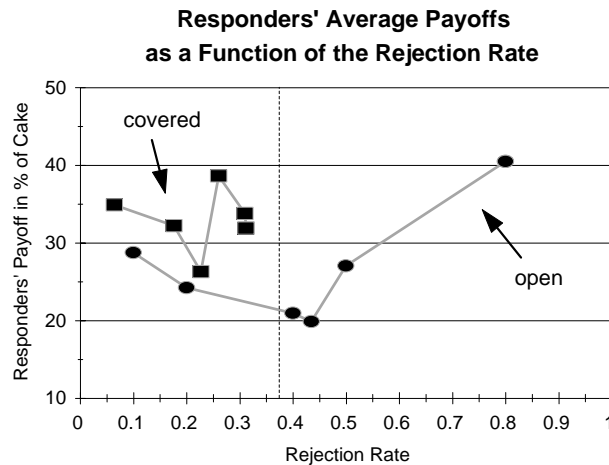


Figure 5.

are exactly indifferent between the equal and the unequal offer. Rejection rates close to that line are least profitable for the responder populations. The very low and especially the very high rejection rates are preferable for payoff maximizing responder populations. However, it should be noticed that the data basis is too small to provide clear statistical evidence for this effect.

### 5.5. Do contributions to group reputation increase efficiency?

Efficiency in our context can simply be understood as the total payoff gained by responders and proposers in a session together. Obviously every rejection reduces the total payoff in a round. The efficiency loss, in points, is exactly the absolute overall number of rejections in a session, multiplied by the cake size of 10. On the other hand, if the proposers are successfully educated by high early rejection rates to increase the number of equal offers, the total number of rejections of a session may be small, since the absolute number of unequal offers is reduced. Hence, if the effect of early punishment is strong enough, then high rejection rates in the beginning of a session need not necessarily lead to less efficiency than in a session with a constant moderate rate of unequal offers and rejections.

In Figure 6, average proposer payoffs are depicted on the  $x$ -axis and the average responder payoffs on the  $y$ -axis. The dots

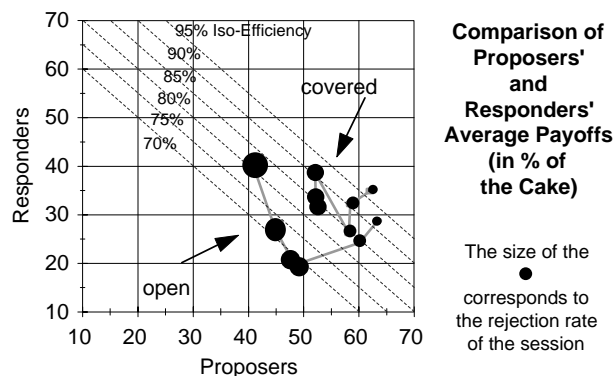


Figure 6.

mark the average responder and proposer payoffs in the sessions, where the size of the dots correspond to the average rejection rate in that session. The diagonal lines are iso-efficiency lines. Different points on the same iso-efficiency line represent allocations with the same efficiency, but a different distribution between proposers and responders. It can be seen that not only the average responder payoff, but also efficiency is lowest with intermediate rejection rates. The highest efficiency is achieved in the session with the lowest rejection rate, where, evidently, most of the efficiency gain is extracted by the proposers. Thus, high rejection rates have a similar effect on efficiency as on responder payoffs. If rejection rates are intermediate, then the contributions to group reputation are too small to outweigh the efficiency losses caused by the rejections. However, we must once again point out that the data basis is small and no conclusive inference can be drawn.

**Observation 6.** The highest responder payoffs and the highest efficiency are observed in the sessions in which rejection rates are either very low or very high.

Finally, in Figure 6, it can immediately be noticed that efficiency in the covered treatment is greater than in the open treatment. This is due to the lower rejection rates coinciding with higher rates of equal offers. The difference is significant at  $p = 0.017$  (one-tailed) according to the Mann–Whitney U-test applied to the average per round total payoff in the single independent sessions.

## 6. SUMMARY AND CONCLUSIONS

We conducted an experiment using a simple ultimatum game. Our two-treatment design enabled us to separate a base rate of rejections motivated by responders' *resistance to unfairness* from an additional rate of rejections that seems to have been motivated by *group reputation* building. We find no support for the hypothesis that rejections are motivated by the *enjoyment of overt punishment*.

In our open response treatment, the responder's choice was reported to the proposer, while it was not reported in the covered response treatment. Rejection rates in the covered response treatment – with an average of 23.3% – were considerable. Since rejections in the covered response treatment can neither have been motivated by individual nor group reputation building, we conclude that the high rejection rates in these sessions are due to the responders' inherent *resistance to unfairness*. This result is completely in line with the central assumption of fairness utility models, such as those suggested by Bolton (1991), Rabin (1993), Dufwenberg and Kirchsteiger (1998), Fehr and Schmidt (1999), and Bolton and Ockenfels (2000).

The significant difference in rejection behavior across treatments, however, is not compatible with any of the mentioned fairness utility models. We observed a significantly higher rate of rejections in the open response treatment than in the covered response treatment. Therefore, we conclude that rejections in the open response treatment are not only motivated by responders' resistance to unfairness, but also by some other motive.

One possible additional motive for rejections in the open treatment can be that responders are contributing to their group's reputation for being "tough." In the context of our experiment, in which each responder met each proposer only once, building an individually profitable reputation was impossible. But, by building a group reputation for rejecting unfair offers, the group of responders attempted and managed to influence the behavior of the group of proposers in the direction of more equal offers.

A different explanation for the higher rejection rates in the open response treatment is that overt and immediate punishment is more enjoyable than veiled and delayed punishment. Since the proposer making an unfair offer was immediately informed of the responders reaction in the open response treatment, such an enjoyment of overt punishment should have led to a sustained higher rate of rejections than in the covered treatment.

The dynamics of rejection behavior provided more evidence for our conclusion. The average per round rejection rates in the open response treatment decreased in the last rounds of each experimental session. In the very last rounds, these rates actually approached the average rejection rate of the covered response treatment, which stayed around 23% and exhibited no time trend. Since the enjoyment of overt punishment should be expected to be stable across rounds, but the value of group reputation declines as the experiment proceeds, we take our results to be supportive of the group reputation hypothesis.

Group reputation in one sense is effective. When facing a group of “tough” responders, proposers tend to make equal offers significantly more often than when facing “compliant” responders. But, although proposers react to punishment as expected, only extreme rejection behavior actually leads to higher profits for the responders. In other words, responders fare best, either if they are markedly compliant or if they are very strict in rejecting unequal offers. The responder groups that only occasionally reject unfair offers, end up paying more for these rejections than they gain from the positive reactions of the proposers. Thus, rejecting unequal offers is always effective, but is not always profitable or efficient. Building a “tough” group reputation only pays, if enough unequal offers are rejected early on.

Finally, we observe that significantly more equal offers are made by proposers in the open than in the covered treatment. Our conjecture is that proposers in early rounds of the open treatment “test” responders’ reactions to unequal offers. In the covered treatment, such testing is not possible. This seems to explain the significantly higher number of unequal first round offers in the open compared to the covered response treatment.

#### ACKNOWLEDGEMENTS

The authors thank the teams of the *RatioLab* at the Hebrew University of Jerusalem and the *Laboratorium für experimentelle Wirtschaftsforschung* at the University of Bonn for their



aid in collecting the data. Furthermore, the authors thank Gary Bornstein, Colin Camerer, Bernd Irlenbusch, Bettina Rockenbach, Reinhard Selten, as well as the seminar and conference audiences in Bari, Berlin, Bilbao, Chicago, Haifa, Meißen, and Paris for many helpful comments. Support by the *Deutsche Forschungsgemeinschaft* through the *Sonderforschungsbereich 303*, by the *German-Israeli Foundation (GIF)*, by the European Union through the TMR program ENDEAR (FMRX-CT98-0238), and by the *Land Nordrhein-Westfalen* is gratefully acknowledged.

#### APPENDIX A. THE WRITTEN INSTRUCTIONS

##### *Player Types:*

There are two types in the experiment: players 1 and 2. After the introduction, each participant draws one of 16 cards.

The drawn card defines the terminal number of the participant.

The terminal number determines the participant's type for the whole experiment.

##### *Structure:*

The experiment consists of eight rounds.

In each round eight pairs of participants are formed: each pair with one player 1 and one player 2.

In every round, every player 1 meets a different player 2, and vice versa.

Thus, no participant meets the same participant a second time.

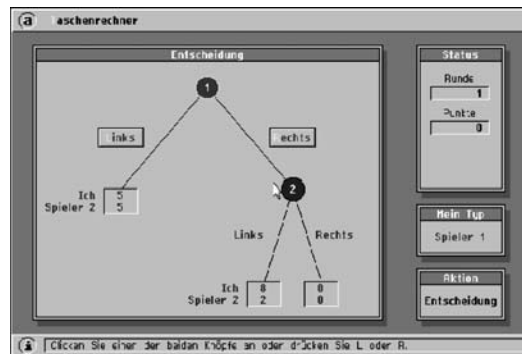
The participants are not allowed to speak with each other during the experiment.

##### *Decisions:*

Each round begins with player 1 choosing one of two alternatives: Left or Right.

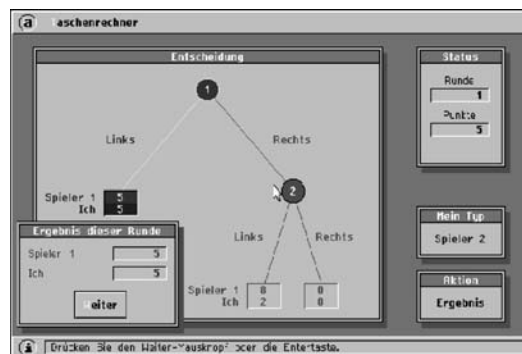
### The Decision Screen for Player 1

Player 1 chooses one of his alternatives by clicking the corresponding mousebutton on the screen, or by pressing the corresponding key on the key board ("L" for "Left", "R" for "Right", resp.).

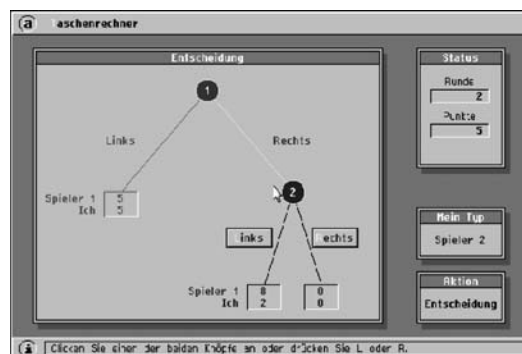


### The Decision Screens for Player 2

If player 1 has chosen "Left", then player 2 is informed about player 1's choice, but has no own decision to make.



If player 1 has chosen "Right", then player 2 chooses one of his alternatives by clicking the corresponding mousebutton on the screen, or by pressing the corresponding key on the keyboard ("L" for "Left", "R" for "Right", resp.).



Player 2 is informed about the choice of player 1.

If player 1 chooses Left, the round ends.

If player 1 chooses Right, player 2 chooses one of two alternatives: Left or Right; player 1 is not informed about the choice of player 2; then the round ends.

*Profits in Points:*

Player 1	Player 2	Player 1 receives	Player 2 receives
Chooses Left	Has no choice	5	5
Chooses Right	Chooses Left	8	2
Chooses Right	Chooses Right	0	0

*The Question Mark:*

If player 1 chooses Right, he will only see a question mark on the screen: since he will not be informed about the choice of player 2, he will neither know his own profit, nor the profit of player 2.

After the last round of the experiment all participants will be informed of their total profits, but the players 1 will not be informed of the value of each received separately.

*Exchange Rate:*

Each point earned in the experiment is equivalent to 50 Pfennigs/75 Agorot.

## APPENDIX B: THE DATA

Session 1: Covered, Bonn									
1	2L 16	11R 15R	13R 3L	8L 1	14R 9L	17L 6	12R 5L	18L 7	
2	2R 15L	11L 3	13R 1R	8R 9R	14R 6L	17L 5	12R 7L	18L 16	
3	2L 3	11R 1R	13R 9R	8L 6	14L 5	17R 7L	12R 16L	18L 15	
4	2L 1	11L 9	13R 6L	8L 5	14R 7L	17L 16	12R 15R	18L 3	
5	2L 9	11R 6L	13R 5L	8R 7L	14L 16	17L 15	12L 3	18L 1	
6	2R 6L	11L 5	13R 7L	8R 16L	14R 15R	17R 3L	12L 1	18L 9	
7	2L 5	11L 7	13L 16	8L 15	14L 3	17L 1	12L 9	18L 6	
8	2R 7L	11R 16L	13L 15	8R 3L	14R 1R	17R 9R	12L 6	18L 5	
Session 2: Covered, Bonn									
1	2L 16	11L 15	13L 3	8L 1	14L 9	17L 6	12L 5	18L 7	
2	2L 15	11L 3	13R 1L	8L 9	14L 6	17L 5	12L 7	18L 16	
3	2L 3	11R 1R	13R 9L	8R 6L	14L 5	17R 7L	12R 16L	18R 15L	
4	2R 1R	11L 9	13L 6	8R 5R	14L 7	17L 16	12L 15	18L 3	
5	2L 9	11R 6L	13R 5R	8R 7R	14R 16L	17L 15	12R 3R	18L 1	
6	2L 6	11L 5	13L 7	8L 16	14R 15L	17L 3	12L 1	18R 9L	
7	2L 5	11L 7	13R 16L	8R 15L	14R 3R	17R 1L	12L 9	18L 6	
8	2L 7	11R 16L	13R 15L	8L 3	14R 1L	17L 9	12L 6	18L 5	
Session 3: Covered, Bonn									
1	2R 16L	11L 15	13L 3	8L 1	14R 9L	17L 6	12L 5	18L 7	
2	2R 15L	11R 3L	13R 1L	8L 9	14R 6L	17L 5	12L 7	18R 16L	
3	2R 3L	11R 1L	13L 9	8R 6R	14R 5R	17L 7	12R 16R	18R 15L	
4	2R 1L	11R 9L	13L 6	8L 5	14R 7L	17L 16	12L 15	18L 3	
5	2R 9L	11R 6L	13R 5L	8R 7L	14R 16L	17L 15	12L 3	18R 1L	
6	2L 6	11R 5R	13R 7L	8L 16	14L 15	17L 3	12L 1	18L 9	
7	2R 5R	11R 7L	13L 16	8L 15	14L 3	17R 1L	12L 9	18L 6	
8	2R 7L	11R 16L	13R 15L	8R 3L	14R 1L	17L 9	12R 6R	18R 5L	
Session 4: Covered, Jerusalem									
1	15R 4L	11L 2	13L 3	9R 1L	14R 8L	16L 6	12R 5L	10L 7	
2	15L 2	11R 3L	13L 1	9R 8R	14R 6L	16R 5L	12L 7	10R 4L	
3	15R 3L	11L 1	13R 8L	9R 6L	14R 5R	16R 7L	12L 4	10L 2	
4	15L 1	11R 8L	13R 6L	9R 5R	14R 7R	16R 4L	12R 2L	10L 3	
5	15R 8R	11L 6	13R 5L	9R 7L	14R 4L	16L 2	12R 3L	10R 1R	
6	15R 6L	11R 5L	13R 7L	9R 4L	14R 2L	16R 3L	12R 1R	10L 8	
7	15L 5	11L 7	13R 4L	9R 2L	14R 3L	16R 1R	12R 8R	10L 6	
8	15L 7	11R 4L	13R 2L	9R 3L	14R 1R	16R 8L	12R 6L	10L 5	
Session 5: Covered, Jerusalem									
1	15L 4	11L 2	13R 3L	9R 1L	14L 8	16L 6	12L 5	10L 7	
2	15R 2R	11R 3R	13R 1L	9R 8L	14R 6L	16L 5	12R 7L	10L 4	
3	15L 3	11L 1	13R 8R	9L 6	14R 5L	16R 7L	12L 4	10L 2	
4	15R 1L	11R 8L	13R 6L	9R 5L	14R 7R	16L 4	12R 2R	10L 3	
5	15R 8L	11R 6L	13R 5L	9R 7L	14L 4	16R 2R	12L 3	10L 1	
6	15L 6	11L 5	13R 7R	9L 4	14R 2L	16L 3	12R 1L	10L 8	
7	15R 5L	11R 7L	13R 4R	9L 2	14R 3L	16L 1	12L 8	10L 6	
8	15R 7R	11R 4R	13R 2L	9L 3	14L 1	16L 8	12L 6	10L 5	
Session 6: Covered, Jerusalem									
1	15L 4	11L 2	13R 3L	9L 1	14L 8	16R 6R	12R 5L	10L 7	
2	15R 2L	11R 3L	13R 1L	9L 8	14R 6L	16L 5	12R 7L	10L 4	
3	15L 3	11L 1	13R 8L	9L 6	14R 5L	16R 7R	12R 4L	10L 2	
4	15R 1L	11R 8L	13R 6L	9L 5	14L 7	16L 4	12R 2L	10L 3	
5	15L 8	11R 6L	13R 5L	9L 7	14L 4	16R 2L	12R 3L	10L 1	
6	15R 6L	11L 5	13R 7L	9L 4	14L 2	16R 3L	12R 1L	10L 8	
7	15L 5	11L 7	13R 4L	9L 2	14R 3L	16R 1L	12R 8L	10L 6	
8	15L 7	11L 4	13R 2L	9L 3	14R 1L	16L 8	12R 6L	10L 5	

Appendix B. Each line represents one round of the session. Each column stands for one match. For example, “15 R 4 L“ reads: terminal 15, proposer, chose Right, terminal 4, responder, chose Left.

## Appendix B. (Continued).

Session 7: Open, Bonn										
1	2L 16	11R 15L	13L 3	8L 1	14R 9L	17R 6L	12R 5R	18R 7L		
2	2R 15R	11R 3R	13R 1R	8R 9L	14R 6L	17R 5R	12R 7L	18L 16		
3	2L 3	11R 1L	13L 9	8R 6L	14R 5R	17R 7L	12R 16R	18R 15R		
4	2R 1L	11R 9L	13R 6L	8R 5R	14R 7L	17R 16R	12R 15R	18L 3		
5	2L 9	11R 6L	13R 5R	8R 7L	14R 16R	17L 15	12R 3R	18R 1L		
6	2R 6L	11R 5L	13L 7	8R 16R	14R 15L	17R 3R	12L 1	18R 9L		
7	2L 5	11R 7L	13L 16	8R 15L	14R 3R	17L 1	12L 9	18R 6L		
8	2R 7L	11R 16R	13L 15	8R 3R	14R 1L	17L 9	12L 6	18R 5R		

Session 8: Open, Bonn										
1	2R 16L	11R 15L	13R 3L	8L 1	14L 9	17L 6	12R 5L	18L 7		
2	2R 15L	11R 3R	13R 1L	8L 9	14R 6L	17R 5L	12R 7L	18L 16		
3	2R 3L	11R 1L	13R 9L	8R 6L	14R 5R	17L 7	12R 16R	18L 15		
4	2R 1L	11R 9L	13R 6L	8R 5L	14L 7	17R 16L	12R 15R	18L 3		
5	2R 9L	11R 6R	13R 5R	8R 7L	14R 16L	17L 15	12R 3L	18L 1		
6	2R 6R	11R 5L	13R 7L	8R 16R	14R 15L	17R 3R	12R 1L	18R 9L		
7	2R 5L	11R 7L	13R 16R	8R 15L	14R 3L	17L 1	12R 9L	18R 6L		
8	2R 7L	11R 16L	13R 15L	8R 3L	14R 1L	17L 9	12R 6L	18R 5L		

Session 9: Open, Bonn										
1	2R 16L	11R 15L	13L 3	8R 1L	14L 9	17R 6R	12R 5L	18R 7R		
2	2R 15L	11L 3	13R 1L	8R 9L	14R 6R	17R 5L	12R 7R	18R 16L		
3	2R 3R	11R 1L	13R 9L	8R 6R	14L 5	17R 7R	12R 16R	18L 15		
4	2R 1L	11R 9R	13R 6R	8R 5L	14R 7R	17R 16L	12L 15	18L 3		
5	2R 9L	11L 6	13R 5L	8R 7R	14L 16	17R 15L	12R 3L	18L 1		
6	2R 6R	11R 5L	13R 7R	8R 16R	14L 15	17R 3R	12R 1L	18L 9		
7	2R 5L	11R 7R	13R 16L	8R 15L	14R 3L	17R 1L	12R 9L	18R 6R		
8	2R 7R	11L 16	13L 15	8R 3L	14R 1L	17R 9L	12R 6R	18R 5L		

Session 10: Open, Jerusalem										
1	15R 4L	11L 2	13L 3	9R 1L	14L 8	16R 6R	12R 5L	10L 7		
2	15R 2R	11R 3L	13R 1L	9R 8R	14L 6	16R 5L	12L 7	10R 4R		
3	15R 3L	11R 1R	13R 8L	9R 6L	14R 5R	16R 7L	12R 4R	10L 2		
4	15R 1L	11R 8R	13R 6R	9L 5	14R 7R	16R 4R	12R 2R	10L 3		
5	15R 8R	11L 6	13L 5	9L 7	14L 4	16L 2	12L 3	10R 1L		
6	15R 6L	11L 5	13R 7L	9R 4L	14L 2	16R 3R	12L 1	10L 8		
7	15R 5R	11L 7	13R 4R	9R 2R	14L 3	16L 1	12R 8L	10R 6L		
8	15R 7L	11L 4	13L 2	9R 3L	14L 1	16L 8	12L 6	10R 5R		

Session 11: Open, Jerusalem										
1	15L 4	11L 2	13R 3R	9R 1R	14L 8	16L 6	12L 5	10L 7		
2	15R 2L	11R 3R	13L 1	9L 8	14L 6	16R 5R	12R 7R	10L 4		
3	15R 3R	11L 1	13L 8	9R 6R	14L 5	16L 7	12L 4	10L 2		
4	15R 1R	11L 8	13R 6L	9L 5	14R 7R	16L 4	12L 2	10L 3		
5	15L 8	11L 6	13R 5R	9R 7R	14L 4	16L 2	12L 3	10L 1		
6	15L 6	11L 5	13L 7	9L 4	14L 2	16L 3	12L 1	10L 8		
7	15R 5R	11L 7	13R 4L	9L 2	14L 3	16L 1	12L 8	10L 6		
8	15L 7	11L 4	13L 2	9L 3	14L 1	16L 8	12L 6	10L 5		

Session 12: Open, Jerusalem										
1	15L 4	11L 2	13R 3L	9R 1R	14L 8	16R 6L	12L 5	10L 7		
2	15R 2L	11R 3L	13R 1L	9L 8	14R 6L	16R 5L	12L 7	10R 4L		
3	15R 3L	11R 1R	13R 8L	9L 6	14R 5L	16R 7R	12R 4L	10R 2L		
4	15R 1L	11L 8	13R 6R	9R 5L	14L 7	16R 4L	12R 2L	10R 3L		
5	15R 8L	11R 6L	13R 5L	9R 7L	14L 4	16R 2L	12R 3L	10R 1L		
6	15L 6	11R 5L	13R 7L	9L 4	14L 2	16R 3L	12R 1R	10R 8L		
7	15R 5L	11R 7L	13R 4L	9R 2L	14R 3L	16R 1L	12R 8L	10R 6L		
8	15R 7L	11R 4L	13R 2L	9R 3L	14R 1R	16R 8L	12R 6L	10R 5L		

## NOTES

1. One of the authors, for example, has a record of repeatedly spending the time and the effort needed to provide this public good. Of course, this is not the only real-life example for punishment of this kind. There are instances

- of furious motorists destroying speed cameras they have detected. Even though the camera does not anymore pose a threat to the motorist once he knows its location, the act of destruction reduces the authorities' inflow of fines to the benefit of fellow motorists. Disappointed customers often provide time and effort to make their negative experience anonymously public on the internet, although they have no apparent benefit from this.
2. Surveys on ultimatum game experiments can be found in Güth (1995), Roth (1995), Camerer and Thaler (1995), or Camerer (2003, chapter 2).
  3. See, e.g., Bolton and Ockenfels' analysis of sequential versus simultaneous prisoners' dilemma games (that lead to very different predictions).
  4. However, there are a number of papers that look at the dynamic aspect of the issue using learning models. For example, Roth and Erev (1995) show that an adaptive learning model (based purely on the adaptation of choices to past outcomes) can produce simulation paths that exhibit some of the features of ultimatum game outcomes without reference to fairness utilities and/or (explicit) reputation. In the long run, however, the similarities between simulation paths and observed behavior are lost. Furthermore, Abbink et al. (2001) find that simple adaptive learning models cannot completely account for responder motivations. They show that many responders not only care about their own payoffs, but also about the payoff received by the proposer. To incorporate more sophisticated behavioral structures, adaptive learning models have been enhanced in a number of ways, e.g. Erev and Roth (1998) Camerer and Ho, (1999) and Camerer et al. (2002).
  5. Reputation effects have been show to be decisive to outcomes in quite a number of experiments, for example: the *borrower-lender game* studied by Camerer and Weigelt (1988), the simple signaling game studied by Brandts and Holt (1992), the *centipede game* studied by McKelvey and Palfrey (1992), and the *chain-store game* studied by Jung et al. (1994).
  6. Gneezy and Stoler (1998) show that in a one-shot situation, social punishment behavior can be observed next to personal revenge.
  7. Along similar lines, Winter and Zamir (1996) suggest that observed ultimatum game behavior may be driven by some form of *population rationality*.
  8. For a survey see Ledyard (1995).
  9. See, for example, Dawes et al. (1988).
  10. Effects related to group reputation effects may be present in other experimental games as well, e.g. in public goods games played with random matching ("strangers" condition). In this setting it is very difficult to build up individual reputation as partners change every round, but contributions can still enhance the group's reputation for being cooperative ("strangers" versus "partners" – fixed matching – comparisons are reported by Andreoni (1988), Weimann (1994), Palfrey and Prisbey (1990), Croson (1996), Burlando and Hey (1997) and Fehr and

- Gächter (2000)). Abbink (2003) conducts a bribery experiment under both partners and strangers conditions. The typical finding is that cooperation is lower among strangers than among partners, but still present. GÄCHTER and FALK (1999) conduct gift exchange game experiments with a fixed versus a never-meet-again matching and find that individual reputation alone cannot entirely explain cooperation.
11. If group reputation plays a role in public good experiments with “strangers” matching, observed contributions should decline towards the end as the incentives for group reputation building go down.
  12. When we started with this project we had not considered this motive. The notion of the enjoyment of overt punishment was put forward by several seminar participants on a number of different occasions. We thank all those who suggested this idea.
  13. Fisher’s two-sample randomization test yields a  $p$ -value of 0.058 (one-tailed).
  14. The figure suggests rising rejection rates in the first three rounds (and a remarkable parallelism between the treatments). However, the figure shows the aggregate data. In the analysis of individual sessions we did not find a strong or significant effect.
  15. This behaviour can be interpreted as an instance of learning direction theory (Selten and Stoecker, 1986; Selten et al., 2001). Switching from the unequal to the equal offer after rejection (i.e., after earning zero) means switching to the option that would have led to a higher payoff of 5. Thus, this switch would be predicted by the theory. After acceptance, which pays 8, no such switch would be expected according to learning direction reasoning.

## REFERENCES

- Abbink, K. (2003), Staff rotation as an anti-corruption policy: An experimental study, *European Journal of Political Economy*, forthcoming.
- Abbink, K., Bolton, G.E., Sadrieh, A. and Tang, F.-F. (2001), Adaptive learning versus punishment in ultimatum bargaining, *Games and Economic Behavior* 37, 1–25.
- Abbink, K. and Sadrieh, A. (1995), RatImage – Research Assistance Toolbox for Computer-Aided Human Behavior Experiments, SFB Discussion Paper B-325, (University of Bonn).
- Bolton, G.E. (1991), A comparative model of bargaining: Theory and evidence, *American Economic Review* 81(5), 1096–1135.
- Bolton, G.E. and Ockenfels, A. (2000), ERC – A theory of equity, reciprocity and competition, *American Economic Review* 90(1), 166–193.
- Bolton, G.E. and Zwick, R. (1995), Anonymity versus punishment in ultimatum bargaining, *Games and Economic Behavior* 10, 95–121.

- Brandts, J. and Holt, C.A. (1992), An experimental test of equilibrium dominance in signaling Games, *American Economic Review* 82, 1350–1365.
- Burlando, R. and Hey, J.D (1997), Do anglo-saxons free-ride more? *Journal of Public Economics* 64, 41– 60.
- Camerer, C. (2003), *Behavioral Game Theory: Experiments on Strategic Interaction*, Princeton University Press, Princeton, NJ.
- Camerer, C. and Ho, T.-H. (1999), Experience-weighted attraction learning in normal form games, *Econometrica* 67, 827–874.
- Camerer, C., Ho, T.-H. and Chong, J.K. (2002), Sophisticated EWA learning and strategic teaching in repeated games, *Journal of Economic Theory* 104, 137–188.
- Camerer, C. and Thaler, R.H. (1995), Anomalies – ultimatums, dictators and manners, *Journal of Economic Perspectives* 9(2), 209–219.
- Camerer, C. and Weigelt, K. (1988), Experimental tests of a sequential equilibrium reputation model, *Econometrica* 56, 1–36.
- Croson, R.T.A. (1996), Partners and strangers revisited, *Economics Letters* 53, 25–32.
- Dawes, R.M., van de Kragt, A.J.C. and Orbell, J.M. (1988), Not me or thee but we: the importance of group identity in eliciting cooperation in dilemma situations: Experimental manipulations, *Acta Psychologica* 68, 83–97.
- Dufwenberg, M. and Kirchsteiger, G. 1998, A Theory of Sequential Reciprocity. *Games and Economic Behavior*, forthcoming.
- Erev, I. and Roth, A.E. (1998), Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria, *American Economic Review* 88(4), 848–882.
- Fehr, E. and Schmidt, K.M. 1999, A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics* 114, 817–868.
- Gächter, S. and A. Falk. (1999), Reputation or Reciprocity? Working paper, (University of Zürich).
- Gale, J., Binmore, K.G. and Samuelson, L. (1995), Learning to be imperfect: the ultimatum game, *Games and Economic Behavior* 8, 56–90.
- Gneezy, U. and Stoler, A. (1998), Why punish? norms and revenge in an experimental game, Working paper, forthcoming in I. Garcia-Jurado, F. Patrone and S. Tijs (eds.), *Game Practice*.
- Güth, W. (1995), On ultimatum bargaining experiments – A personal review, *Journal of Economic Behavior and Organization* 27, 329–344.
- Güth, W., Schmittberger, R. and Schwarze, B. (1982), An experimental analysis of ultimatum bargaining, *Journal of Economic Behavior and Organization* 3, 367–388.
- Jung, Y.J., Kagel, J.H. and Levin, D. 1994, On the existence of predatory pricing: An experimental study of reputation and entry deterrence in the chain-store game, *RAND Journal of Economics* 25, 72–93.



- Keser, C. and van Winden, F. (2000), Conditional cooperation and voluntary contributions to public goods, *Scandinavian Journal of Economics* 102, 23–39.
- Ledyard, J.H. (1995), Public goods: A survey of experimental research, in Kagel J. and Roth, A.E. (eds), *Handbook of Experimental Economics*. (Princeton University Press, Princeton, NJ), 111–194.
- McKelvey, R.D., and Palfrey, T.R. (1992), An experimental study of the centipede game, *Econometrica* 60, 803–836.
- Palfrey, T.R. and Prisbey, J.E. (1996), Altruism, reputation and noise in linear public goods experiments, *Journal of Public Economics* 61, pp. 409–427.
- Rabin, M. (1993), Incorporating fairness into game theory and economics, *American Economic Review* 83, 1281–1302.
- Roth, A.E. (1995), Bargaining experiments, in Kagel, J. and Roth, A.E. (eds), *Handbook of Experimental Economics* (Princeton University Press, Princeton).
- Roth, A.E., Prasnikar, V. Okuno-Fujiwara, M. and Zamir, S. (1991) Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo, *American Economic Review* 81, 1068–1095.
- Roth, A.E. and Erev, I. (1995), Learning in extensive form games: Experimental data and simple dynamic models in intermediate Term, *Games and Economic Behavior* 8, 164–212.
- Thaler, R. H. (1988), The ultimatum game, *Journal of Economic Perspectives* 2, 195–206.
- Weimann, J. (1994), Individual behaviour in a free riding experiment, *Journal of Public Economics* 54, 185–200.
- Winter, E. and Zamir, S. (1996), Experimenting Ultimatum Bargaining with Virtual Players, Working Paper (Hebrew University Press, Jerusalem).

Klaus Abbink, School of Economics, The University of Nottingham,  
University Park, Nottingham NG7 2RD United Kingdom.

Tel: 44-115-9514768; Fax: 44-115-9514159.

E-mail: klaus.abbink@nottingham.ac.uk;

Abdolkarim Sadrieh, Department of Economics and Center for Economic Research, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands.

Tel: 31-13-4662938; fax: 31-13-4663042.

E-mail: sadrieh@uvt.nl;

Shmuel Zamir, Center for Rationality at the Hebrew University Jerusalem, Israel and EUREQua – Paris 1 and Crest—LEI/ENPC, 28 rue des Saints Pères, 75007 Paris, France.

Tel : 33-1-44 58 27 70; fax: 33-1-44 58 27 72

E-mail: zamir@ensae.fr;