Inequity Aversion and Individual Behavior in Public Good Games: An Experimental Investigation

Astrid Dannenberg[a], Thomas Riechmann[b], Bodo Sturm[a] and Carsten Vogt[c]

[a]Centre for European Economic Research (ZEW)
[b]Faculty of Economics and Management, University of Magdeburg
[c]Department of Business Administration, Leipzig University of Applied Sciences

E-mail:
dannenberg@zew.de, thomas.riechmann@ww.uni-magdeburg.de,
sturm@zew.de, vogt@wiwi.htwk-leipzig.de

Preliminary version – do not quote
February 2007

Abstract
We present a simple two-steps procedure for a within-subject tests of the inequity aversion model of Fehr and Schmidt (1999). In the first step, subjects played modified ultimatum and dictator games and were classified according to their preferences. In the second step, subjects with specific preferences according to the Fehr and Schmidt model were matched into pairs and interacted with each other in a standard public good game and a public good game with punishment possibility. Our results show that the specific composition of groups significantly influences the subjects' performance in the public good games. We identify the aversion against advantageous inequity and the information about the co-player's type as the main influencing factors for the behavior of subjects.

JEL classification: C91, C92, H41
Keywords: individual preferences, inequity aversion, experimental economics, public goods
Acknowledgements: Financial support from the German Science Foundation is gratefully acknowledged.

1 Introduction

Within experimental economics there is a growing number of stylized facts which contradict the model of rational payoff maximizing actors. People cooperate in social dilemmas such as public good games (Ledyard 1995), they reject high amounts of money in the ultimatum game (Güth et al. 1982, Camerer 2003) and last but not least they make positive contributions in the dictator game (Kahneman et al. 1986, Forsythe et al. 1994, Camerer 2003). The contradiction between the standard economic model of selfish behavior and empirical observations has been a challenge for both theorists and experimentalists. In the last ten years a number of theories that try to answer this question has been developed. Most of these theories are based on the assumption that people have some kind of other-regarding, or social, preferences. These approaches seek to overcome the discrepancies between standard game-theoretical prediction and experimental observation by altering the underlying utility function of the subjects, but stick to the assumption that subjects behave rationally. The models by Bolton and Ockenfels (2000) and Fehr and Schmidt (1999) are prominent examples for this approach. They assume that people are willing to pay money in order to avoid unequal payoff distributions. Besides these "inequity-aversion" theories other approaches focus on intentions of subjects, i.e. the way a subject behaves affect whether a player cares positively or negatively about that subject. Rabin (1993) is the pioneering paper in this direction, while Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006) extended Rabin's approach to extensive form games. A common property of models with other-regarding preferences is that subjects are heterogeneous in their preferences. This implicates that theoretical prognoses about individual behavior may differ between subjects for the same decision problem.

In this study we focus on the model of inequity aversion by Fehr and Schmidt (1999), in the following F&S. There are two reasons for doing this. Firstly, the F&S model is able to explain an impressive amount of experimental evidence not in line with the standard model of selfish behavior. Secondly, F&S use a model which is from a theoretical point of view quite parsimonious as only two additional parameters are added to the individual utility function. Moreover, both parameters of the model can be estimated with the help of simple laboratory techniques.

One interesting implication of models with other-regarding preferences such as F&S is that they allow within-subject tests, i.e. controlled experiments with the same subject but different decision problems. Due to the fact that these theories predict – given different preferences –

different behavior of subjects, one may test hypotheses at the individual level with the following two-steps procedure. In a first step, individual other-regarding preferences are measured by means of appropriately designed games. In a second step, the same subjects interact with each other in a controlled environment under specific rules for which hypotheses regarding the individual behavior have been derived in advance. Under the assumption that preferences are stable at least within a short time period, this approach allows a robust test of such models in the laboratory. Remarkably, this approach has already been mentioned by Fehr and Schmidt (1999), p. 847,

> *“One of the most interesting tests of our theory would be to do several different experiments with the same group of subjects. Our model predicts a cross-situation correlation in behavior. For example, the observations from one experiment could be used to estimate the parameters of the utility function of each individual. It would then be possible to test whether this individual’s behavior in other games is consistent with his estimated utility function.”*

Our study implements such a two-steps procedure. In the first step, we measure the individual F&S preferences by means of two simple experiments, a modified ultimatum game and a modified dictator game. In the second step, subjects with specific preferences are matched together into pairs and interact with each other in a standard public good game and a public good game with a punishment possibility. We distinguish between three groups of pairs. In particular, we form a group of “fair” subjects where both players of the pair are highly inequity averse, a group of “egoistic” subjects where both players are very little inequity averse, and a “mixed” group where one player is “fair” and the other one is “egoistic”. Due to the composition of treatments with subjects with specific preferences we are able to derive and test hypotheses according to the F&S model.

Our results differ from the results described in comparable papers. Firstly, the weight of aversion against disadvantageous inequity varies very little throughout our subject pool and has a median of zero. Secondly, our results show that the specific composition of groups significantly influences the subjects' performance in the public good games: As long as subjects are informed about the type of their opponent, “fair” groups contribute more to the public good than “egoistic” or “mixed” groups. It turns out, furthermore, that explicit information is a key factor for this difference in behavior: As long as “fair” subjects are not informed on the fact that their co-player is “fair”, too, they act like “egoistic” subjects. Only the explicit information that they are playing with a “fair” co-player significantly enhances their contributions. Uninformed “fair” players are not significantly more cooperative than “egoistic” ones.

The remainder of the paper is organized as follows. Section 2 sets the stage by describing the F&S model which underlies our experiment. Section 3 describes the design of our experiment including treatments and hypotheses. Section 4 presents the experimental results. Section 5 summarized and discusses our results. Section 6 concludes.

2 Theoretical background: The model of Fehr and Schmidt (1999)

2.1 Preferences

According to Fehr and Schmidt (1999) individuals are not exclusively motivated by the absolute payoff they can earn but also value allocations due to their distributional consequences. Particularly, assuming that individuals suffer from inequality F&S introduce the following utility function for subject i:

$$U_i(\pi_i, \pi_j) = \pi_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i}^{n} \max\{\pi_j - \pi_i, 0\} - \beta_i \frac{1}{n-1} \sum_{j \neq i}^{n} \max\{\pi_i - \pi_j, 0\} \qquad (1)$$

where π_i and π_j denote the absolute payoffs to subjects i and j, respectively, n denotes the total number of players involved in some decision problem, $\alpha_i \geq 0$ measures the impact of i's disutility from disadvantageous inequality while $\beta_i \geq 0$ measures the corresponding impact of advantageous inequality. In the two player case which is particularly relevant for our experimental setting, (1) reduces to

$$U_i(\pi_i, \pi_j) = \pi_i - \alpha_i \max\{\pi_j - \pi_i, 0\} - \beta_i \max\{\pi_i - \pi_j, 0\}. \qquad (2)$$

F&S assume $\beta_i < 1$, i.e. players are not willing to „burn“ their money to eliminate advantageous inequality. In addition, they assume that players put a stronger weight on disadvantageous inequality, i.e. $\alpha_i \geq \beta_i$. In our experiment, we will obtain the weights α_i and β_i from a modified ultimatum game and a modified dictator game (see section 3.1).

2.2 Voluntary contribution games

2.2.1 The standard voluntary contribution game

The assumption of such preferences may have a strong impact on the theoretical predictions on the outcomes in several classes of games. In a public good game for example, preferences of the F&S-type may lead to much higher cooperation rates compared to the predictions derived by standard economic theory. To see this, look at the following voluntary contribution game. Each player $i = 1,...,n$ is given some initial endowment y which can be devoted to the production of some public good. Player i`s contribution to the public good is denoted by g_i, the production function for the public good is simply given by the sum over all contributions $\sum_{j=1}^{n} g_j$. Let us assume that the marginal per capita return of an investment in the public project is given as some constant $1/n < a < 1$. Then the monetary payoff for player i is given by $\pi_i(g_1,...,g_n) = y - g_i + a\sum_{j=1}^{n} g_j$. Obviously, this game constitutes a social dilemma. The marginal return to an investment in the public good is a while the marginal costs for such an investment amount to 1. Thus, for player i it is a dominant strategy to choose $g_i = 0$. Since this holds for all players identically, the unique equilibrium of this game is characterized by contributions $g_j = 0 \; \forall j$ and the public good will not be provided at all. However, the provision would be beneficial since the collective marginal return is na which is clearly above the marginal costs of provision. Hence, the social optimum is achieved if each player contributes his entire initial endowment to the public good leading to payoff $\pi_i^{SO} = any$ which is above the payoff players receive in the Nash equilibrium ($\pi_i^{NE} = y$).

F&S have shown that this result is fundamentally altered if players are endowed with inequality aversion according to (1). They prove the following results:

1. If $a + \beta_i < 1$, then it is a dominant strategy for player i to choose $g_i = 0$.
2. Let k, $0 \leq k \leq n$, denote the number of players with $a + \beta_i < 1$. Then, if $k/(n-1) > a/2$, there exists a unique equilibrium with $g_i = 0 \; \forall i \in \{1,...,n\}$.
3. If $k/(n-1) < (a + \beta_j - 1)/(\alpha_j + \beta_j)$ for all players $j \in \{1,...,n\}$ with $a + \beta_j > 1$, then equilibria with positive contributions to the public good exist. All k players with $a + \beta_i < 1$ choose $g_i = 0$ while all other players contribute $g_j = g \in [0, y]$.

The intuition behind these results is not too difficult. Firstly, if a player with $a+\beta_i<1$ invests one monetary unit in the public good his monetary return is a while he gains a maximum non-monetary utility of β_i. Now, if the sum of both returns is less than one it is obviously the best strategy not to invest into the public good, irrespectively of what other players do. Secondly, if there are sufficiently many players with $a+\beta_j<1$, then player i will not be willing to contribute even if he shows stronger inequality aversion, i.e. for him $a+\beta_i>1$ holds. The reason is that relatively few "fair" players are not able to sufficiently reduce disadvantageous inequality. Thirdly, if there are sufficiently many players with $a+\beta_j>1$, they can sustain cooperation amongst themselves, "*even if the other players do not contribute. However, this requires that the contributors are not too upset about the disadvantageous inequality toward the free riders.*" (Fehr and Schmidt 1999, p. 840).

2.2.2 The voluntary contribution game with punishment

The idea that punishment of defective players may increase contribution rates to the public good is straightforward. In a setting with standard preferences, however, punishment is a non credible threat. Imagine a two-stage game: Stage one is the voluntary contribution game as described in the section above. Stage two of the game incorporates the possibility for players to enact some punishment on their opponents. Since punishment is costly it will not be carried out by rational players interested only in their absolute material payoff on the second stage. Since players anticipate the outcome on the second stage they will defect in the first stage of the game.

This outcome is substantially altered if preferences of the F&S-type are involved. Fehr and Schmidt show that the existence of a group of so called "conditionally cooperative enforcers" may enhance the prospects for cooperation. These individuals must show sufficiently strong aversion against advantageous inequality, i.e. their preferences must obey $\beta_i \geq 1-a$. In addition, punishment must not be too costly. Let c denote the costs of punishment. If

$$c<\frac{\alpha_i}{(n-1)(1+\alpha_i)-(n'-1)(\alpha_i+\beta_i)}\ \forall i \in [1,...,n'] \tag{3}$$

and all other players $i \in [n'+1,...,n]$ do not care about inequality, i.e. for them $\alpha_i=\beta_i=0$, then the following strategies form a subgame perfect equilibrium: In the first stage each player contributes $g_i = g \in [0,...,y]$. If each player does so, there are no punishments on the second stage of the game. If, however one of the players $i \in [n'+1,...,n]$ deviates and chooses $g_i < g$, then each enforcer carries out some punishment $p_{ji} = (g - g_i)/(n'-c)$ while all other players do not punish. If one of the enforcers deviates by choosing $g_i < g$, or if any other player chooses $g_i > g$, then one Nash-equilibrium of the punishment game is being played.

2.2.3 Introducing uncertainty

The analysis in F&S is based upon the assumption that players know their opponents' type, i.e. they know the number of players k with preferences $\beta_i < 1 - a$. For this reason, in most of our experimental treatments, we informed the participants, previous to the public good games (see section 3.1), on how their opponent had behaved in the modified ultimatum and dictator game played before. Thus, these subjects were principally able to derive the corresponding type of their co-player. In one treatment, we did not inform the subjects about their opponent. These subjects were only able to predict their opponent's type with some probability.[1] In this case, a "fair" player with preference $\beta_i \geq 1 - a$ will choose to contribute to the provision of the public good if the following condition is met:

$$\frac{a + \beta_i - 1}{\alpha_i + \beta_i} \geq \frac{E(k)}{n-1},$$

where $E(k)$ denotes the expected value of k. In the case of two players this condition can be rearranged as follows:

$$\beta_i \geq \frac{E(k)}{1 - E(k)} \alpha_i + \frac{1}{1 - E(k)} (1 - a). \tag{4}$$

Note that as a consequence of introducing uncertainty the parameter α_i matters. This is different in the case of perfect knowledge which can be easily seen by setting $E(k) = 0$. Then, (4) reduces to $\beta_i \geq 1 - a$. Obviously, for increasing values of α_i it becomes more difficult to fulfil (4), i.e. in order to meet the condition players must have higher values of β_i. The intuition behind this is as follows. On the one hand, if a "fair" but uniformed subject contributes to the public good he runs the risk of having an "egoistic" opponent and,

---

[1] In the following we assume risk neutral behavior.

therefore, being exploited. On the other hand, if he does not contribute, he runs the risk of having a “fair” opponent and possibly exploiting her. In other words, positive contributions to the public good would increase the risk of disadvantageous inequity and decrease the risk of advantageous inequity. Hence, the subject is only willing to contribute, if his aversion against disadvantageous inequity is sufficiently low and his aversion against advantageous inequity is sufficiently large. Since the condition $\beta_i \geq 1-a$ for informed subjects is more easily met than the condition (4) for uninformed subjects, we should observe a higher level of cooperation from informed players than from uninformed players.

3 Experimental design

3.1 Games

We used four different games (game A, B, C and D) in our experimental design. The design of these games is presented in the following.[2] Game A – which resembles the responder’s decision problem in the ultimatum game[3] – is designed to measure the subjects’ aversion against disadvantageous inequity. Each subject has to decide in 22 cases (numbered from #1 to #22) in the role of player 1 between two pairs of payoffs (pair I and pair II) each with an amount of money for herself and another subject in the role of player 2. Payoffs (see the left part of Table 1) are chosen in a way that – except for #1 – subjects always have to choose between “pair I”, a disadvantageously unequal division of 10.00 €, and “pair II”, an equal distribution with 2.00 € for both players. All cases were arranged in a descending order by the amount of money subjects could earn in pair I. In this game, a purely selfish subject should choose pair I from #1 to #20 and pair II for #21 and #22.[4] A subject strongly disliking disadvantageous inequity, in contrast, would choose pair I in #1 and pair II from #2 to #22. Subjects with other-regarding preferences according to F&S between these two extremes would be expected to switch from choosing pair I to pair II after #2 but prior to #21.

We describe individual behavior in game A as consistent if (1.) a subject has a unique switching point from pair I to pair II and (2.) the switching point is between #2 and #21. Regarding the first condition, a subject with aversion against disadvantageous inequity

---

[2] See the appendix for the instructions that we distributed to our participants.

[3] The difference to the original ultimatum game is the fact that the conflict point payoffs (in €) are changed to (2, 2) instead of the original (0, 0). See section 4.1 for the discussion of another potential difference between the decision problem in game A and the original ultimatum game.

[4] In the following we assume rational behavior of all subjects.

consistent with the F&S model who switches for a specific case from pair I to pair II should choose for all subsequent cases pair II. As the payoffs for player 1 in pair I are arranged in descending order, a switch back to pair I in any of the subsequent decisions is not consistent. This would lead to a lower own payment and to higher disadvantageous inequity than in the case that was rejected before. In relation to the second condition, it is useful to consider the decision cases outside of the "consistent area" between #2 and #21. A subject who chooses pair II in #1 already is not regarded as consistent because she could attain an equal allocation with higher own payoff by choosing pair I. A subject who chooses pair II in #22 only or does never switch to pair II at all has a negative value for the weight of aversion against disadvantageous inequity ($\alpha_i < 0$), i.e. likes disadvantageous inequity, and is therefore not consistent with the F&S model. With the subject's switching point we can determine the upper and lower bounds of the individual α_i. We approximate the individual value for α_i by choosing the mean of the corresponding interval (see Table 1). There are two exceptions to this rule. Firstly, we cannot determine an upper bound for α_i of a subject who switches from pair I to pair II in #2. Therefore, we assign to those subjects the value of the lower bound, $\alpha_i = 2.18$. Secondly, we assign the value $\alpha_i = 0$ to a subject who switches from pair I to pair II in #21, although the corresponding interval for this case is $-0.08 \leq \alpha_i \leq 0.04$. In doing so, we might include subjects with a negative value for α_i but otherwise we would run the risk of omitting subjects with a non-negative value of α_i from the analysis.

Game B – which resembles the decision problem in the dictator game – is designed to measure the subjects' aversion against advantageous inequity. Again, each subject had to decide between two pairs of payoffs (pair I and pair II) each with an amount of money for herself in the role of player 1 and another subject in the role of player 2 in 22 cases (from #1 to #22; see the right part of Table 1). Payoffs are chosen in a way that – except in #21 and #22 – subjects had to choose between "pair I", an extremely unequal but advantageous distribution of 10.00 €, and "pair II", an equal distribution of different amounts from 0.00 to 21.00 €. All cases were arranged in an ascending order by the amount of money subjects could earn in pair II. In this game, a purely selfish subject would choose pair I from #1 through #20 and pair II for #22. In the case of #21, this subject would be indifferent between pair I and pair II. A subject strongly disliking advantageous inequity would always choose pair II. Subjects with "fairness preferences" according to F&S would be expected to switch from choosing pair I to pair II after #1 but before #21.

We label individual behavior in game B as consistent if (1.) a subject has a unique switching point from pair I to pair II and (2.) this switching point is between #2 and #22., i.e. if the individual weight of aversion against advantageous inequity meets $0 \le \beta_i < 1$. Relating to the first condition, a subject with aversion against advantageous inequity consistent with the F&S model switching from pair I to pair II at one point should also choose pair II in all cases after the switching point. As the payoffs for player 1 in pair II are ordered in an ascending order, a switch back to pair I in any of the subsequent cases is not consistent. This would lead to the same advantageous inequity than was rejected before but now with higher opportunity costs in terms of equal payoffs for both players. For the second condition, we consider again the decision cases outside of the "consistent area" between #2 and #22. A subject choosing pair II already in #1 has $\beta_i \ge 1$, i.e. is willing to "burn" money in order to produce equal payoffs. A subject who does not switch at all displays affection for advantageous inequity. Both behavioral patterns are not consistent with F&S.

Table 1: Payoffs in game A and game B

		game A					game B				
		pair I		pair II			pair I		pair II		
		payoffs (in €) for player					payoffs (in €) for player				
	#	1	2	1	2	α_i	1	2	1	2	β_i
switching point from pair I to pair II	1	5.00	5.00	2.00	2.00	-	10.00	0.00	0.00	0.00	1.00
	2	4.44	5.56	2.00	2.00	2.18	10.00	0.00	0.50	0.50	0.98
	3	4.42	5.58	2.00	2.00	2.13	10.00	0.00	1.00	1.00	0.93
	4	4.39	5.61	2.00	2.00	2.02	10.00	0.00	1.50	1.50	0.88
	5	4.36	5.64	2.00	2.00	1.90	10.00	0.00	2.00	2.00	0.83
	6	4.32	5.68	2.00	2.00	1.77	10.00	0.00	2.50	2.50	0.78
	7	4.29	5.71	2.00	2.00	1.66	10.00	0.00	3.00	3.00	0.73
	8	4.24	5.76	2.00	2.00	1.54	10.00	0.00	3.50	3.50	0.68
	9	4.19	5.81	2.00	2.00	1.41	10.00	0.00	4.00	4.00	0.63
	10	4.14	5.86	2.00	2.00	1.30	10.00	0.00	4.50	4.50	0.58
	11	4.07	5.93	2.00	2.00	1.18	10.00	0.00	5.00	5.00	0.53
	12	3.92	6.08	2.00	2.00	1.00	10.00	0.00	5.50	5.50	0.48
	13	3.86	6.14	2.00	2.00	0.85	10.00	0.00	6.00	6.00	0.43
	14	3.81	6.19	2.00	2.00	0.79	10.00	0.00	6.50	6.50	0.38
	15	3.68	6.32	2.00	2.00	0.70	10.00	0.00	7.00	7.00	0.33
	16	3.53	6.47	2.00	2.00	0.58	10.00	0.00	7.50	7.50	0.28
	17	3.33	6.67	2.00	2.00	0.46	10.00	0.00	8.00	8.00	0.23
	18	2.85	7.15	2.00	2.00	0.30	10.00	0.00	8.50	8.50	0.18
	19	2.72	7.28	2.00	2.00	0.18	10.00	0.00	9.00	9.00	0.13
	20	2.22	7.78	2.00	2.00	0.10	10.00	0.00	9.50	9.50	0.08
	21	1.43	8.57	2.00	2.00	0.00	10.00	0.00	10.00	10.00	0.03
	22	0.10	9.90	2.00	2.00	-0.14	10.00	0.00	10.50	10.50	0.00

Similar to game A, we can determine the upper and lower bounds for the individual's β_i with a subject's switching point. We approximate the individual value for β_i by choosing the mean of the corresponding interval (see Table 1). As before, there is an exception to this rule. We assign the value $\beta_i = 0$ to a subject who switches from pair I to pair II in #22, although the corresponding interval in this case is $-0.05 \leq \beta_i \leq 0$. In doing so, we might include subjects with a negative value for β_i but otherwise we would possibly omit subjects from the analysis with a non-negative value for β_i .

Game C is a standard two-player public-good (PG) game with a voluntary contribution mechanism. Two players get a fixed balance of 3.00 € for show-up and are endowed with 10.00 € each. They decide simultaneously how much (if any) money from the endowment to contribute to a public good. Each monetary unit that the subject keeps for herself raises her payoff by exactly that amount. Both subjects receive 0.70 € for each 1.00 € contributed to the public good, i.e. the marginal per capita return (MPCR) is constant and equal to 0.7. The game was played over 10 rounds with the number of rounds as common knowledge. After each round subjects were informed about the own contribution and the contribution of the co-player as well as the payoffs of both players.

Game D consists of two stages. Stage 1 is equivalent to game C, i.e. subjects play a two-player PG game with a show-up fee of 3.00 €, an endowment of 10.00 € and a constant MPCR of 0.7. Stage 1 in game D is followed by a stage 2. In this stage subjects have the possibility to assign her co-player negative points, i.e. a punishment mechanism is introduced. Each negative point reduces the payoff of the co-player by 1.00 €. However, the assignment of negative points is costly. Each negative point assigned reduces the punisher's own payoff by 0.17 € or 0.50 € (respectively, depending on the treatment). Again, the game was played over 10 rounds with the number of rounds as common knowledge.[5]

Subjects were paid separately for games A and B and games C and D. The payments from game A and B were computed as follows: All subjects within a session were randomly matched into pairs of subjects. After this, it was determined (again by chance) whether game A or B is relevant for the payoffs. After the selection of the relevant game, a random draw selected which number of the payoff list (between #1 and #22) was relevant. Finally, a

---

[5] Theoretically, due to possible punishment and costs of punishment in game D, it is possible to encounter cases with negative payoffs. We were prepared to handle these cases (payments would have set to zero). Luckily, though, a case like this did not occur during the experiments.

random draw decided which person within a pair determined the payoffs, i.e. which subject decides in the role of player 1. According to this rule, each of the 22 decision cases in game A and B had the same chance to be relevant for the payment. Subjects were informed about this payoff rule in advance and we checked the comprehension of this design feature in a quiz before the experiment started. The payments from game C and D were determined in a similar way: After both games, a random draw determined which game (C or D) was relevant. Following this decision, one of the 10 periods was selected randomly and the payments were realized according to the decisions in this round. As before, the payoff rule was common knowledge to all participants.

Subjects played games A and B at first. Subjects who behaved consistently in A and B and fulfill the conditions $1 > \beta_i \geq 0$ and $\alpha_i \geq 0$ were invited to play another two games (C and D).[6]

3.2 Subject pool and treatments

We ran 21 sessions with 18 to 20 subjects in each session. All in all, 414 subjects participated in games A and B and 144 of these subjects were invited to play games C and D. Sessions lasted about 90 minutes and the average earning were 5.90 € for games A and B and 15.10 € for games C and D. The sessions were conducted in November 2006 at the Magdeburg Experimental Laboratory (MaXLab). The experiment was fully computerized and anonymous.[7] In the laboratory, the subjects were randomly allocated to separate cabins and had no mutual contact during or after the experiment. The main characteristics of our subject pool are displayed in Table 2. Though the majority of our subjects are students of economics the fraction of non-economists (38%) is quite high. Almost two third of our subjects have experience in experiments, i.e. have participated in at least one experimental session before.

---

[6] Those subjects who were not invited to take part in games C and D played a one-shot prisoner’s dilemma game (the only purpose of which was to avoid a zero payment for subjects who didn’t play game C and D). After this game these subjects were paid and left the laboratory. The mean payment (3.70 €) was accordingly low. In the following we do not refer to the results of this game.

[7] We used Z-tree for programming. See Fischbacher (1999).

Table 2: Subject pool – descriptive statistics

	absolute frequency	relative frequency in percent
total	414	100.0
consistent choices	319	77.0
study		
Management Science (MS)	153	36.9
Economics (Ec)	34	8.2
other Economics (othEc)	70	16.9
MS or Ec or othEc (Econ)	257	62.1
other than Econ (NonEcon)	157	37.9
experienced (at least one experiment before)	270	65.2
male subjects	233	56.3
female subjects	181	43.7

In order to check the robustness of our design, we implemented two modifications during the experiment. In two of our sessions, we modified the order of play. While in most sessions, subjects first played game A followed by game B, in two sessions, we had subjects play B first, then A. Moreover, in seven of our sessions, payoffs in game A were slightly different: The payoff in pair II of game A was 0.00 € for both subjects instead of 2.00 €. Both changes do not lead to significantly different distributions of the values for α_i and β_i at the 5% level (K-S test, two-tailed).[8] Therefore, we felt free to pool the data of all 21 sessions conducted of games A and B.

In order to analyze the predictive power of the F&S model as well as some comparative static effects we defined several treatments for games C and D, which are described below.[9] Treatments differ with respect to (1.) the preference parameter β_i of the subjects that form a group in games C and D, (2.) the state of information about the co-player's behavior in games A and B and (3.) the value of parameter c, which defines the marginal costs of punishment in game D.

Our four treatments of game C differ with respect to the composition of the (two-person)-groups playing the game. Details are displayed in Table 3. Thereby, one group of two players (i = 1, 2) makes one statistically independent observation. In treatment EGO, two subjects with $\beta_i < 0.3$ are matched together to form a common group. Treatments MIX, FAIR and FAIR(ni) are defined correspondingly. In all treatments except for FAIR(ni), all subjects are

---

[8] Unless it is explicitly noted, in the following all tests are two-tailed.

[9] The definition of treatments is determined by the structure of the distribution of α_i and β_i within our subject pool. Due to the fact that there is virtually no dispersion for α_i we had to focus the definition of treatment on parameter β_i.

informed about their respective co-player's former behavior from games A and B. In FAIR(ni), players did not receive any information about each other's former behavior.

Table 3: Treatments in game C

Treatment	Parameter β_i , i = 1, 2	Information	Observations
EGO	$\beta_i < 0.3$	yes	35
MIX	$\beta_1 < 0.3 \wedge \beta_2 \geq 0.3$	yes	13
FAIR	$\beta_i \geq 0.3$	yes	12
FAIR(ni)	$\beta_i \geq 0.3$	no	12
Σ			72

Note: All subjects in FAIR and FAIR(ni) have $\alpha_i = 0$.

All subjects who played game C also completed game D, the PG game with punishment possibilities. As we distinguish between high costs of punishment and low costs, we have two treatments in game D corresponding to each of the treatments in C, one with high and one with low costs of punishment. The different treatments in game D are displayed in Table 4. As in game C before, in most of the treatments, subjects were informed about their co-players' behavior in games A and B, again with the notable exception of treatment FAIR(h, ni).

Table 4: Treatments in game D

Treatment	Parameter β_i , i = 1, 2	Information	Costs of Punishment	Observations
EGO(l)	$\beta_i < 0.3$	yes	low	9
EGO(h)	$\beta_i < 0.3$	yes	high	26
MIX(h)	$\beta_1 < 0.3 \wedge \beta_2 \geq 0.3$	yes	high	10
FAIR(h)	$\beta_i \geq 0.3$	yes	high	12
FAIR(h, ni)	$\beta_i \geq 0.3$	no	high	12
Σ				69

Notes: (1.) As we have only three independent observations in the MIX(l) treatment it is omitted from hypotheses testing in game D. (2.) None of the subjects with $\beta_i \geq 0.3$ in MIX(h) fulfills condition (3), p. 6. (3.) All subjects in FAIR(h) and FAIR(h, ni) have $\alpha_i = 0$.

3.3 Hypotheses

Based on the definition of treatments and the theoretical considerations in section 2 we are able to derive specific hypotheses which follow from the F&S model for our subject pool. Thereby we focus – at least for the non-parametric tests in section 4 – on the last period of both games C and D. By doing this we can exclude any repeated game effects, which may come into play if there is a repeated interaction between only two subjects.[10] In the following we assume that, whenever F&S predict the existence of multiple equilibria, subjects will play an equilibrium with positive contributions. Therefore, we can derive the following hypotheses regarding the contributions to the public good:

Hypotheses 1 for game C

a) In treatments EGO and MIX, zero contributions of all subjects should be observed.

b) In treatment FAIR, cooperation, i.e. positive contributions to the public good, should be observed.

c) In treatment FAIR, cooperation should be observed more frequently than in FAIR(ni).

Hypotheses 2 for game D[11]

a) In treatments EGO(l), EGO(h) and MIX(h) zero contributions of all subjects should be observed.

b) In treatment FAIR(h) cooperation, i.e. positive contributions to the public good, should be observed.

c) In treatment FAIR(h, ni) cooperation should be observed more frequently than in FAIR(h, ni).

4. Results

The results section consists of two main parts. The first part analyzes the subjects' behavior in games A and B. The second part focuses on games C and D and refers to the treatments and hypotheses described above.

---

[10] See Fehr and Schmidt (1999), p. 843, for a similar reasoning.

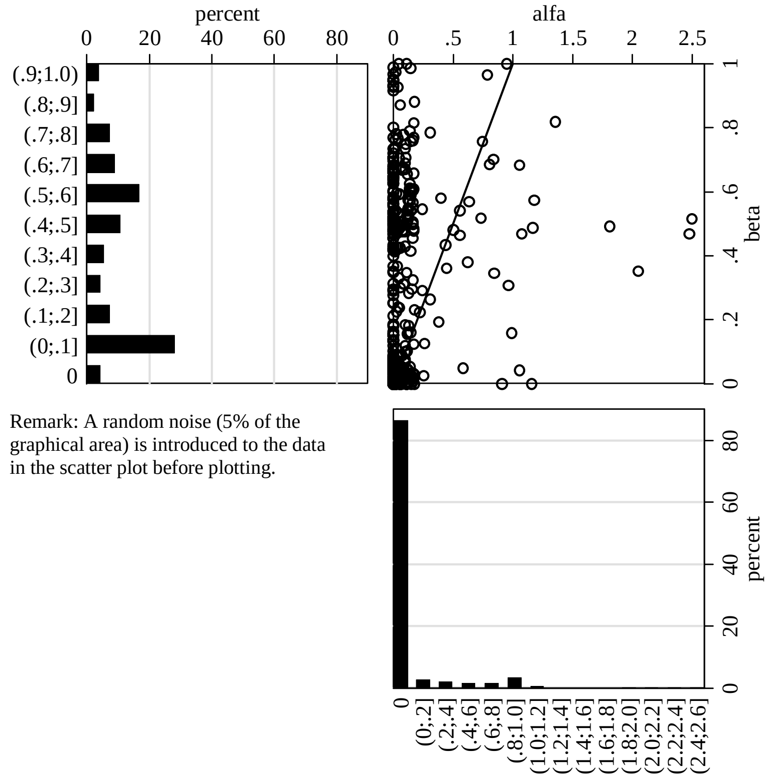
[11] There is no treatment in game D where punishment is a part of an equilibrium strategy. Note that in treatment MIX(h) none of the subjects with $\beta_i \geq 0.3$ fulfills condition (3), p. 6, i.e. there are no subjects who may enforce cooperation. In treatment FAIR(h) cooperation without punishment is an equilibrium because – in the case of mutual cooperation – there are no incentives for subjects with $\beta_i \geq 0.3$ to punish each other. In this case, punishment reduces the own payoff and creates advantageous inequality. Consequently, mutual cooperation without punishment is an equilibrium.

4.1 Behavior in games A and B

As discussed in the previous section, we are able to select subjects with consistent preferences which are in line with the assumptions $\alpha_i \geq 0$ and $0 \leq \beta_i < 1$ in the F&S model. All in all, out of 414 subjects who participated in the experiment 319 (77%) behaved consistently in games A and B. Thereby, we do not detect any significant correlations between the socio-economic characteristics in Table 2 and the consistency of choices at the 5% level (Spearman's ρ).

Figure 1 presents the distribution of the F&S parameters, α_i and β_i. A brief look at the distribution of the values for α_i, the weight for the aversion against disadvantageous inequity, shows that nearly 87% of all subjects behave selfishly in game A (lower right in Figure 1). There is a second peak in the range of $0.8 < \alpha_i \leq 1$. These are subjects who choose the payoff in pair II (2.00 € for both) instead of the payoff in pair I (3.92 € for themselves and 6.08 € for the other subjects) in #12 of game A. The mean value for α_i (0.098) is significantly above zero (t test, p = 0.000). The median value (0), however, indicates that selfish behavior is the dominant pattern in our subject pool. The distribution of the values for β_i, the weight for the aversion against advantageous inequity, looks quite different (upper left in Figure 1). We observe two peaks in the distribution. Firstly, for $0 < \beta_i \leq 0.1$, i.e. for subjects who behave rather selfishly and switch from pair I to pair II in #21 or #22 of game B. Secondly, there is a peak for the range $0.4 < \beta_i \leq 0.6$, i.e. for subjects who switch from pair I to pair II between #10 and #13 and who prefer a rather equal allocation of the 10 € vis-à-vis an advantageous but very unequal allocation. The mean of the β_i values (0.358) is significantly above zero (t test, p = 0.000), the median is equal to 0.425. The scatter plot in Figure 1 (upper right) shows the joint distribution of the two parameters. Remarkably, in our subject pool we found very few subjects (38 of 319, i.e. 12% of consistent choices) meeting F&S' condition $\alpha_i \geq \beta_i$. The corresponding data points lie below the 45-degree line in the scatter plot. It is apparent that α_i and β_i are not significantly correlated and a test for correlation confirms this (Spearman's $\rho = 0.091$, p = 0.103).

Figure 1: Distribution of F&S parameters for consistent choices



Remark: A random noise (5% of the graphical area) is introduced to the data in the scatter plot before plotting.

Given the quite heterogeneous subject pool in our experiment we are able to test whether there are any correlations between the socio-economic characteristics in Table 2 and the individual values for α_i and β_i. We observe a significant negative correlation between subjects who study Management Science, Economics or a related field (e.g. “Wirtschafts-informatik”) and the value for β_i (Spearman’s $\rho = -0.137$, p = 0.015).[12] However, this is the only significant correlation between α_i or β_i and the socio-economic characteristics in Table 2 at the 5% level.

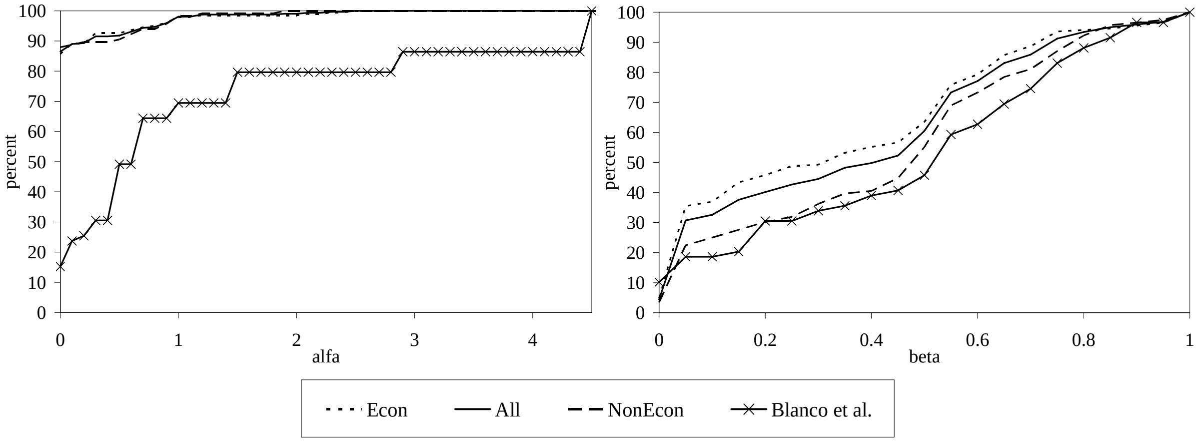
In a recent paper, Blanco et al. (2006) present an experimental test of the F&S model which is similar to our approach. In their experiment, one and the same cohort of subjects plays four different games: (1.) an ultimatum game to derive α_i, (2.) a modified dictator game to derive β_i, (3.) a one-shot PG game and (4.) a sequential prisoner’s dilemma. In order to control for strategic uncertainty the strategy method is used for each game. All of their 72 subjects were non-economists. While the modified dictator game they used is practically identical to game B in this study, our game A differs from the ultimatum game in Blanco et al. (2006) in an important aspect. In the ultimatum game with the strategy method, each subject reacts to a

---

[12] See Marwell and Ames (1981), Frank et al. (1993, 1996), and Carter and Irons (1991) for similar findings.

specific proposal of her co-player, i.e. there is a distinct element of strategic interaction in this game. In our game A, however, each subject herself decides "only" between different distributions of 10.00 € and the equal payoff with 2.00 € for both, which means that there is no direct interaction between both subjects and no room for strategic considerations. In other words, contrary to Blanco et al. (2006) we use two "non-strategic games" for the elicitation of F&S preferences. While this difference does not matter from a theoretical point of view – neither with regard to the F&S model nor the standard game theory – it may influence the behavior of subjects.

Figure 2: Cumulated density for F&S parameters – comparison



Because of the similarity between both approaches of revealing individual preferences according to the F&S model, it is interesting to compare the individual values for α_i and β_i. Figure 2 presents the cumulated densities for α_i and β_i for economists ("Econ"), non economists ("NonEcon"), and all subjects ("All") in the present study (These data are the same as the ones in Figure 1). Additionally the corresponding cumulated densities of the above mentioned Blanco et al. study are depicted.[13] The difference in the cumulated densities for α_i between our subjects and Blanco et al.'s participants is remarkable and highly significant (K-S test, p = 0.010). While in Blanco et al. only 15% have $\alpha_i = 0$, in our study the fraction of subjects without any preference for disadvantageous inequity is more than five times higher. Regarding the weight of advantageous inequity, β_i, the differences between Blanco et al. and our data are not significant at the 5% level (K-S test). On the other hand, as one may expect given the negative correlation between studying economics and β_i, the

---

[13] There are 13 subjects in Blanco et al. (2006) who did not behave consistently according to our definition. Therefore, the parameters of only 59 subjects are displayed for Blanco et al. in Figure 2.

difference in the cumulated density for economists and non economists in the present study is significant at the 5% level (K-S test).

4.2 Behavior in games C and D

The results section for the games C and D consists of five parts. The first and second parts analyze the subjects' contributions in the standard PG game (game C) and the PG game with punishment possibility (game D). Here, we are able to test whether the prognoses from the F&S model apply to our subject pool. The third part investigates the effect of information and the fourth part the effect of punishment. Finally, we present Tobit estimations for the contribution to the public good in game C and the punishment behavior in game D.

4.2.1 Hypotheses testing of the F&S model in game C

A total of 144 subjects played the PG game C. These subjects had behaved consistently in the games A and B and meet the conditions $1 > \beta_i \geq 0$ and $\alpha_i \geq 0$. The mean contribution per period over all subjects and all periods is 6.0 €. A first hint regarding the predictive power of both parameters in the F&S model is Spearman's ρ, the non-parametric coefficient of correlation, between the contribution of subject i to the public good in period t, $C_{i,t}$, and α_i and β_i, respectively. We get $\rho_{C_{i,t},\alpha_i} = -0.113$ (p = 0.000) and $\rho_{C_{i,t},\beta_i} = 0.088$ (p = 0.000) for a total of N = 1440 observations. Thus, there is a weak but highly significant correlation between the contributions and both parameters in the F&S utility function. While subjects with a higher β_i seem to contribute slightly more to the public good, α_i has a small negative effect on contributions.[14]

In order to test our hypotheses derived from the F&S model, we analyze the behavior of the subjects in the treatments EGO, MIX, and FAIR (see Table 3). These treatments differ only in their composition of individual types, "egoistic" or "fair" ones. The attributes "egoistic" and "fair" arise only from the value of β_i, as in the absence of uncertainty about the co-player's type, the subjects' contributions in the standard PG game only depend on β_i.[15] Still, we have

---

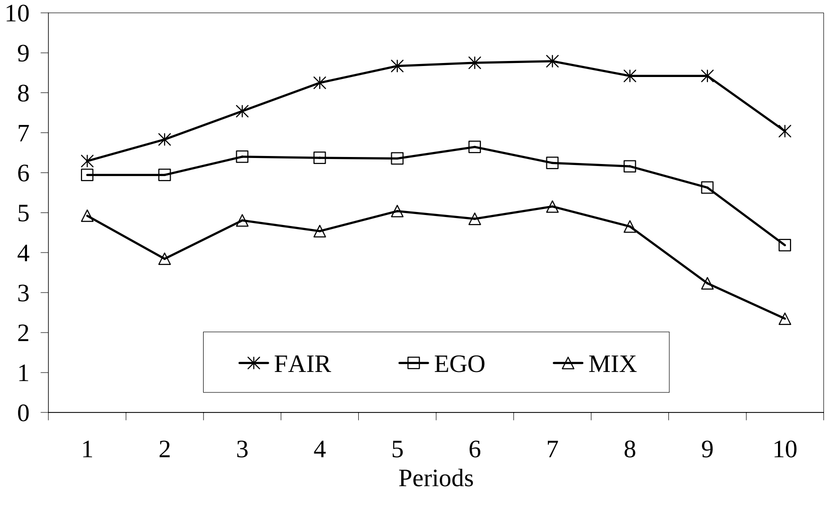
[14] In this case we assume that the direction of causality goes from the parameter of the F&S utility function to the contribution behavior.

[15] The Spearman correlation test shows that α_i has a small negative effect on the contributions. This effect, however, is too small to change the results. If we control for α_i and apply the non-parametric tests to all subjects with $\alpha_i = 0$ (85% of our subjects in game C and D) we find the same results as in the tests including all subjects.

to bear in mind that by "ignoring" the value of α_i in the definition of our types we include into the analysis subjects meeting F&S' condition of $\alpha_i \geq \beta_i$ as well as subjects who do not meet this condition. As most of our subjects have $\alpha_i = 0$ the "egoistic" subjects with low values of β_i often meet the condition whereas the "fair" subjects do very rarely.

The prognosis of the F&S model is the same for the treatments EGO and MIX (see Table 3 for treatments in game C). Contributions of zero are the only equilibrium. This is the same allocation that standard economic theory predicts for both games. In contrast, for the FAIR treatment, according to F&S there exist the "standard" non-contribution equilibrium, and other equilibria with positive contributions. Figure 3 shows the mean contributions in game C for the three treatments with information, namely EGO, MIX, and FAIR. The mean contribution per period is 7.9 € for the FAIR treatment, 6.0 € for EGO, and 4.3 € for MIX. In each period of all treatments the share of subjects contributing a positive amount of money to the public good is significantly greater than zero (Binomial Sign test, one-tailed, p = 0.000). We can state, therefore, that on the one hand the behavior of the subjects in the EGO and the MIX treatment does correspond neither with the F&S model nor with standard theory. On the other hand, behavior in FAIR is in line with the F&S model but not with standard theory.

Figure 3: Contributions in game C



Result 1: The contributions of the subjects in the EGO and the MIX treatment to the public good are significantly higher than the levels predicted by F&S and by the standard model of pure selfishness. The contributions in FAIR are in line with F&S but not with the standard model.

As already explained above (section 3.3), our analysis focuses particularly on the behavior in the final period. 54% of the subjects in the EGO treatment, 58% of the MIX subjects but only 25% of the FAIR subjects "defect" in the final period by contributing nothing to the public good. The shares are not much different if we take contributions below 3 € as "defection".[16] Thus, the shares of defecting subjects are relatively high in the EGO and the MIX treatment compared to the FAIR treatment. In consideration of our hypothesis 1 for game C, this is what we would have expected. In each treatment more than 50% of the subjects behave in line with the F&S predictions. A Mann-Whitney U test (MW U test) shows that the differences in the final period contributions between EGO and MIX are not significant at the 5% level. In contrast, the differences in the final period between MIX and FAIR (p = 0.003) as well as between EGO and FAIR (p = 0.041) are significant. According to this, the F&S model has some explanatory power for the behavior in the final period of the standard PG game. In order to check whether this result is robust, we also apply a χ^2 test to the distribution of the final period contributions. We can reject the hypothesis that cooperation (positive contribution) and defection (zero contribution) is equiprobable for the FAIR treatment (p = 0.004) whereas we cannot reject this hypothesis for EGO and MIX at the 5% level. The same is true if we take contributions below 3 € as defection. Comparing the final period contributions between treatments, we can reject the hypothesis that the share of defecting subjects in the FAIR treatment equals the share in EGO (p = 0.000) or MIX (p = 0.003). We cannot reject this hypothesis for the comparison between EGO and MIX. Again, the level of defection, zero contribution or contribution below 3 €, does not change the result.[17]

> *Result 2: The F&S model has explanatory power for the behavior in the final period of the PG game: Subjects in the FAIR treatment contribute more to the public good than subjects in the EGO and the MIX treatment. Furthermore, the share of subjects who "defect" in the final period is lower in the FAIR treatment than in the EGO and the MIX treatment.*

Figure 3 illustrates the relatively poor performance of the subjects in the MIX treatment. Even though the differences between MIX and EGO are not significant, it is remarkable that the contributions of the MIX treatment are always lower. The MIX treatment is the only

---

[16] If we take contributions below 5 € as "defection", 54% of the subjects in the EGO treatment, 73% of the MIX subjects, and 29% of the FAIR subjects defect in the final period.

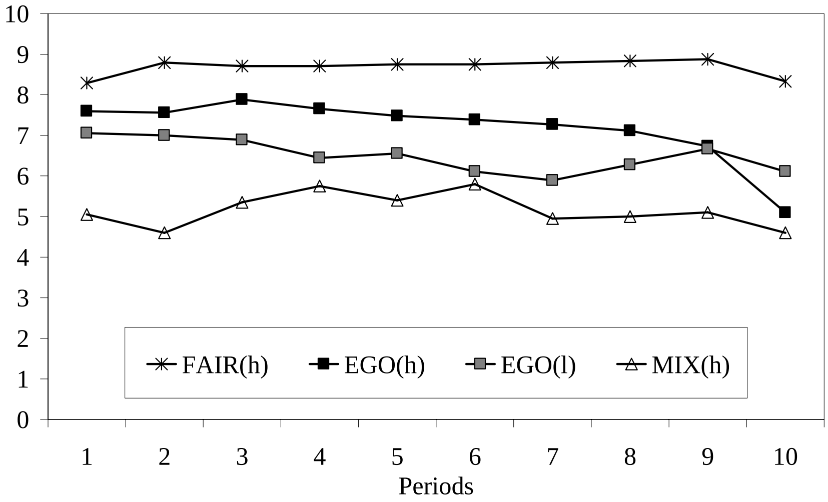
[17] If we interpret contributions below 5 € as "defection", the difference in the shares of defecting subjects between FAIR and EGO as well as between FAIR and MIX are still significant at the 5% level. In this case, the difference between EGO and MIX is significant, too.

heterogeneous treatment where two different types of subjects make up a pair, namely a “fair” subject and an “egoistic” subject. It is interesting to find out whether both types choose low contributions or whether one of the types is responsible for this development. The more detailed analysis indicates that the first period plays an important role where the only information subjects have about their co-player is the behavior in games A and B. The “egoistic” subjects in the MIX treatment contribute on average 6.3 € in the first period whereas the “fair” subjects contribute only 3.5 €. Remember that the mean per-period contribution over all treatments and all periods is 6.0 €. This difference is weakly significant (MW U test, p = 0.057). After the first period, the contributions quickly converge to a relatively low level so that we do not find any more significant differences between both types. Considering the share of defecting subjects in the first period, 38% of the “fair” subjects chose zero contributions compared to only 8% defecting “egoistic” subjects. The difference between the share of defecting subjects is significant (χ^2 test, p = 0.024). We conclude that it is the knowledge that the co-player is an “egoistic” type that prevents the “fair” subjects from exerting high contributions in the PG game. We will return to this aspect later.

4.2.2 Hypotheses testing of the F&S model in game D

In order to test our hypotheses for game D we consider the behavior of the subjects in the treatments EGO(h), MIX(H), and FAIR(h) (see Table 4). The F&S prognoses for these three treatments are the same as for game C. Zero contributions constitute the only equilibrium for the treatments EGO(h) and MIX(h). Theoretically, in case of the FAIR(h) treatment equilibria with positive contributions are feasible. Figure 4 shows the development of the mean per-period contributions over time for all three treatments. The mean contribution over all periods for the FAIR(h) treatment is 8.7 €, for EGO(h) 7.2 €, and for MIX(h) 5.2 €. In all periods and in all treatments, the share of cooperating subjects is significantly greater than zero (Binomial Sign test, one-tailed, p = 0.000), indicating that the contributions of the subjects in the EGO(h) and MIX(h) treatments are generally higher than the F&S predictions.

Figure 4: Contributions in game D



Result 3: In the EGO(h) and the MIX(h) treatments, the individual contributions to the public good in the PG game with punishment possibility are significantly higher than the levels predicted by F&S. Again, contributions in FAIR(h) are in line with F&S but not with the standard model of pure selfishness.

In game D, the differences in the final period contributions between EGO(h) and FAIR(h) as well as between MIX(h) and FAIR(h) are significant only at the 10% level (MW U test). The differences in the final period contributions between EGO(h) and MIX(h) are insignificant.

48% of all subjects in the EGO(h) treatment and 45% of the MIX(h) subjects contribute nothing in the final period. In contrast to this, only 12% of the FAIR(h) subjects defect (zero contributions). These shares remain almost unchanged if we take contributions below 3 € or below 5 € as defection. As already observed in game C, “fair” subjects seem to be more likely to cooperate in the final period than “egoistic” subjects. By means of a χ^2 test, we analyze this aspect in more detail. The test shows that for the treatments EGO(h) and MIX(h) we cannot reject the hypothesis that defection and cooperation is equiprobable at the 5% level. In contrast, we can reject this hypothesis for the FAIR(h) treatment (p = 0.004). The comparison between the treatments exhibits that the share of defecting subjects does not significantly differ between EGO(h) and MIX(h) but that it does differ between EGO(h) and FAIR(h) (p = 0.000) as well as between MIX(h) and FAIR(h) (p = 0.000). This result is independent from the level of defection, namely zero contributions, contributions below 3 €, or contributions below 5 €.

> *Result 4: The F&S model has some predictive power for the behavior in the final period of the PG game with punishing opportunities, although it is somewhat weaker than in the standard PG game: The FAIR(h) subjects contribute more to the public good than the subjects in the EGO(h) and the MIX(h) treatment. The share of defecting subjects is significantly lower in the FAIR(h) treatment than in the other treatments.*

4.2.3 Effect of information

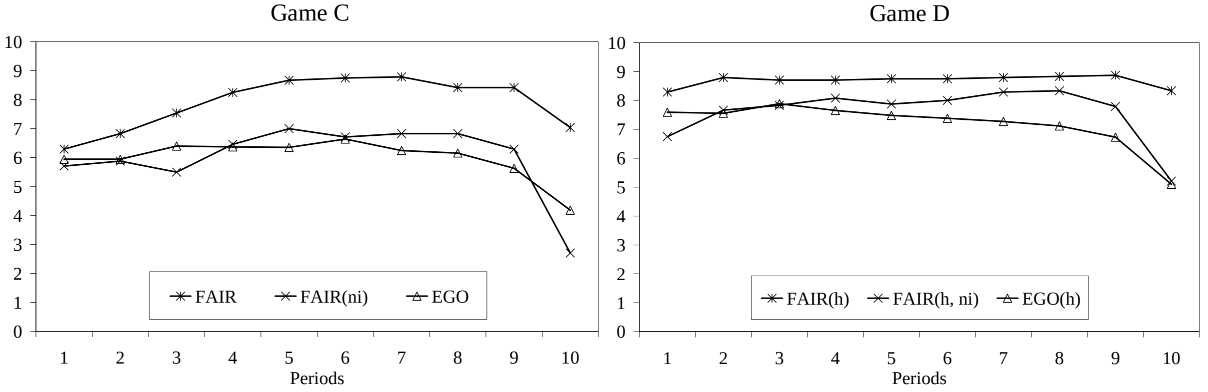
As mentioned above, we varied the amount of information subjects have about the former behavior of their co-player. Except for the treatment FAIR(h, ni) subjects were informed about how their co-player behaved in the games A and B previous to the games C and D. As the treatments FAIR(h) and FAIR(h, ni) differ only with regard to information and not with regard to the F&S parameters α_i and β_i, the comparison between these two treatments enables us to analyze the isolated effect of information on the contributions to the public good.[18]

The mean contributions of the FAIR(h) and the FAIR(h, ni) treatments are shown in Figure 5. For reference, we also insert the contributions of the subjects in the EGO(h) treatment. The contributions of the uninformed subjects are lower than the contributions of the informed subjects in both games and in all periods. For the final period in both games, the differences are significant at the 5% level (MW U test). Remarkably, the information about the co-player seems to play an important role only in the final period, when strategic incentives to cooperate do not exist any longer. In contrast, we do not find significant differences between final period contributions in the FAIR(ni) and the EGO treatment as well as between FAIR(h, ni) and EGO(h). It seems that, in both games, the uniformed FAIR(ni) subjects in the final period behave similar to the EGO subjects. Remember that in both games the differences between FAIR and EGO in the final period are significant at least at the 10% level.

---

[18] All 48 subjects in treatments FAIR(h) and FAIR(h, ni) have $\alpha_i = 0$ and $\beta_i \geq 0.3$ (see Table 4).

Figure 5: Contributions – effect of information



Let us consider the behavior of subjects in the final period in more detail. In this period of game C, 71% of the FAIR(ni) subjects defect whereas in the FAIR treatment only 25% defect. These shares are almost the same for all three levels of defection, zero contributions, contributions below 3 €, and contributions below 5 €. Thus, the share of defecting subjects in the FAIR(ni) treatment is even larger than the share in the EGO treatment where 54% defect in the final period. A χ^2 test shows, however, that the shares of defecting subjects between FAIR(ni) and EGO is not significant at the 5% level whereas the difference between FAIR(ni) and FAIR is highly significant (p = 0.000). The analysis of the final period of game D confirms this result, although the finding is somewhat weaker as it depends on the level of defection. If we equate defection with zero contributions, 8% of subjects in the FAIR(h) treatment, 17% of the FAIR(h, ni) subjects, and 38% of the EGO(h) subjects defect in the last period. The differences are not significant at the 5% level, neither between FAIR(h) and FAIR(h, ni) nor between EGO(h) and FAIR(h, ni). If, instead, we take contributions below 3 € (below 5 €) as defection, the shares of defecting subjects do not change in the treatments FAIR(h) and EGO(h) but for the FAIR(h, ni) treatment the share increases to 33% (42%). In these cases the differences between FAIR(h) and FAIR(h, ni) are significant at the 5% level (χ^2 test) whereas the differences between FAIR(h, ni) and EGO(h) are not significant, again indicating that the behavior of the uniformed FAIR(h, ni) subjects is similar to the subjects in the EGO(h) treatment.

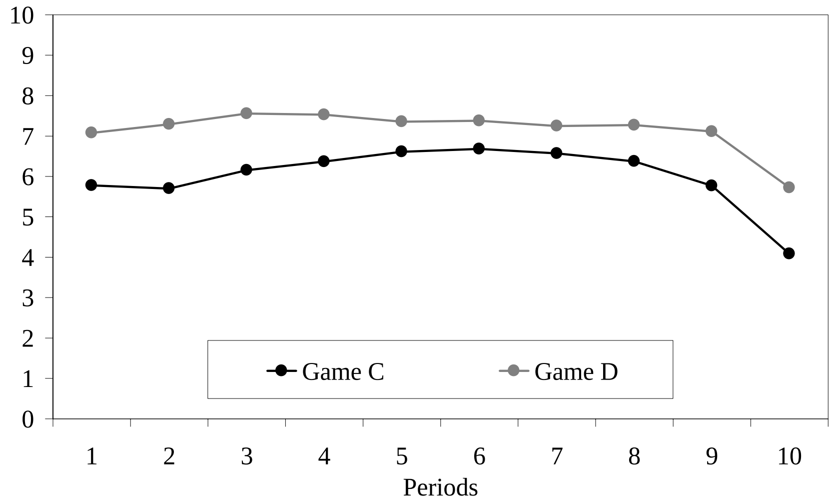
The finding that uninformed FAIR-groups in treatments FAIR(ni) and FAIR(h, ni) perform very similar to EGO-groups in EGO and EGO(h), while informed FAIR-groups in treatments FAIR and FAIR(h) contribute significantly more seems to allow for a conclusion regarding subjects' expectations. As long as subjects do not have information about their co-players, they act as if playing against an EGO-type. It takes explicit information that the co-player is a "fair" one to cause "fair" behavior. Thus, subjects' "default" expectation is that their co-player is an EGO-type.

> *Result 5: In the final period of both games, uninformed subjects contribute significantly less than informed subjects.*

4.2.4 Effect of punishment

The comparison between the games C and D illustrates the influence of punishment possibilities on the contributions to the public good. According to the standard model of purely selfish behavior, the introduction of punishment possibilities does not change the prediction for the contributions. As punishment is costly, a rational individual would not punish and the dominant strategy is still to contribute nothing to the public good. Experimental tests (Sefton et al. 2002, Fehr and Gächter 2000, 2002) show, however, that the introduction of punishment usually increase the contributions. This is also what we observe in game D. Figure 6 illustrates the mean per-period contributions over all subjects for game C and game D.

Figure 6: Contributions – effect of punishment



The mean per-period contribution over all periods is 6.0 € in game C and 7.2 € in game D. It seems that punishment has a positive effect on the contributions. A Wilcoxon Matched-Pairs Signed-Ranks test over all subjects who played C and D confirms that the difference between both games is highly significant (p = 0.000).

The use of two different prices of punishment allows us to analyze the effect of punishment costs on the contributions.[19] Therefore, we compare the mean contributions between the treatments EGO(l) and EGO(h) (see Figure 4). The differences are not significant at the 5% level (MW U test). Thus, the costs of punishment do not seem to have any influence on the level of contributions to the public good.

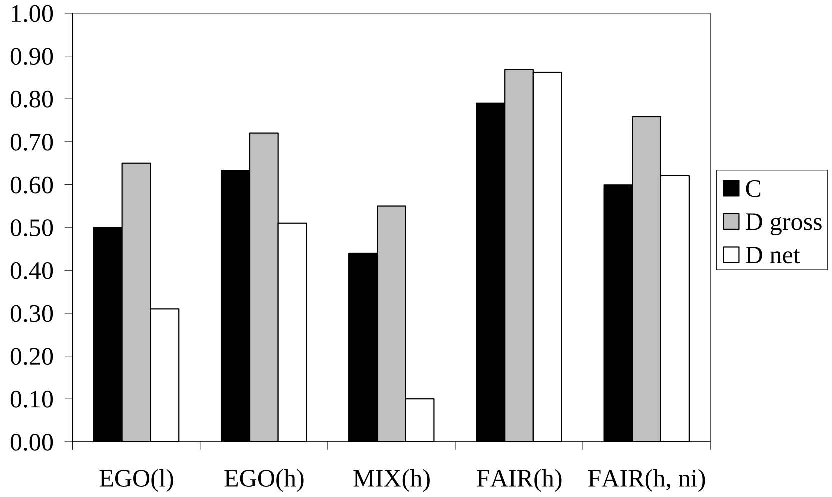
> *Result 6: The introduction of an option to punish the co-player significantly increases the mean contribution to the public good. The costs of punishment do not have an effect on the level of contributions.*

The possibility of punishing the co-player may increase or decrease the payoffs of the subjects. The payoffs will increase if higher contributions outweigh the costs of punishment. They will be lower if the costs of punishment outweigh the increase in contributions. In order to measure the change of payoffs we use a measure of efficiency defined as $Eff = \left(\pi - \pi^{NE}\right) / \left(\pi^{SO} - \pi^{NE}\right)$ with π giving the actual payoff, π^{NE} giving the payoff in the Nash equilibrium, and π^{SO} giving the payoff in the social optimum. In game D, we can furthermore distinguish between gross efficiency and net efficiency. The concept of gross efficiency incorporates the notion that payoffs solely result from contributions. This concept neglects the fact that individual payoffs may be reduced by punishment received as well as by costs of punishment imposed to others. Gross efficiency in the punishment condition is usually larger than efficiency in the no-punishment condition because punishment tends to increase contributions and, by that, (gross) payoffs (see Figure 6). Net efficiency contains the actual payoffs that include both contribution and punishment. Net efficiency in game D can be larger or smaller than efficiency in game C, depending on which of the two effects, change of contributions or punishment, dominates. Note that net efficiency can even be negative because there is no punishment in the Nash equilibrium. If, for example, two subjects contribute nothing but punish each other at the same time, their payoffs will be lower than the payoff in the Nash equilibrium. Consequently, the net efficiency would be negative in this case. Figure 8 shows the degree of efficiency in the games C and D for all treatments.

---

[19] See the Tobit estimation in the next section for an analysis of the marginal costs of punishment on the punishment behavior.

Figure 8: Efficiency



As expected, gross efficiency in game D always exceeds the efficiency in game C. Net efficiency, however, is lower than efficiency in game C for treatments EGO(l), EGO(h) and MIX(h). This is particularly compelling in the treatment MIX(h): The difference between gains from contributions only and these gains net of costs induced by punishment received and imposed is highest in the MIX-groups, i.e. groups consisting of different types of subjects, one “egoistic” and one “fair” type. Homogeneous groups, i.e. groups of “fair” or “egoistic” types only, perform better than mixed groups. One reason for the poor performance of MIX-groups is the fact that members of these groups tend to react to punishment with retaliation instead of increases of contributions. As a consequence, we often witness strong escalation of punishment over time without any effect on contributions.[20]

The comparison between net efficiency of the treatments shows that the difference between MIX(h) and EGO(h) as well as the difference between MIX(h) and FAIR(h) are significant at the 5% level (MW U test). In contrast to all other treatments, the net efficiency of FAIR(h) and FAIR(h, ni) in the punishment condition is somewhat larger than the efficiency in the no-punishment condition. In these cases punishment is able to slightly increase the efficiency. Only groups of “fair” subjects use the option of punishment to their mutual advantage. At the same time, the positive effect of information is still at work. The comparison between the efficiency of the treatments also shows that the difference between FAIR(h) and FAIR(h, ni) is significant (MW U test, p = 0.011).

> *Result 7: Punishment increases the net efficiency compared to a no-punishment environment only in the FAIR treatments. Information about the co-player*

---

[20] See also the Tobit estimation in the next section for this aspect.

strengthens the positive effect. In the treatments EGO(l), EGO(h), and MIX(h) the net efficiency decreases, this being particularly distinctive in the MIX(h) treatment.

4.2.5 Multivariate analysis

In the sections before we use univariate tests, which cannot account for several factors that could affect contributions or punishment behavior. Therefore, in this section we try to investigate the effects of α_i and β_i on the behavior of all subjects in games C and D with a multivariate analysis. In the following we focus on the variables "contribution of subject i in period t" for game C and "punishment points from subject i to j in period t" for game D. Due to the fact that both dependent variables are censored from below and from above, Table 5 presents Tobit estimates.

In game C the contribution of the co-player from the previous period has a strong and significantly positive effect on the contribution of subject i in the current period. In contrast to our expectations and to the intuition from the F&S model, the parameter α_i has a significant and negative effect on the contributions of subject i over all treatments.[21] On the other hand, except for treatment FAIR(ni), parameter β_i has a positive effect on i's contribution, which is significant at the 1% level. This is indicated by the corresponding interaction dummies, i.e. the product of the treatment dummies and the parameter β_i. The positive effect of β_i on the contributions is stronger for EGO than for MIX and FAIR, i.e. β_i seems to have a stronger effect on the contributions if the parameter itself is small. Remarkably, the effect of β_i on contributions in treatment FAIR(ni) is significantly negative. There are no significant effects for the treatment dummies.

Result 8: While parameter α_i has a significantly negative effect on the contributions to the public good over all treatments, the effect of β_i varies over treatment and is significantly positive except for treatment FAIR(ni). The effect of β_i is stronger in treatments with subjects who have small values of β_i.

---

[21] It is not possible to include interaction dummies between treatments and α_i in the estimation, because of the high amount of observations with $\alpha_i = 0$.

Table 5: Tobit estimations

Independent variables	Dependent variable Game C: contribution of subject i in period t	 Game D: punishment points from subject i to j in period t
Constant	-2.3608 (0.4415)***	0.1876 (0.0780)**
Contribution of co-player j in period t		0.0734 (0.0083)***
Contribution of co-player j in period t-1	0.9227 (0.0421)***	
Contribution of subject i in period t		-0.0356 (0.0071)***
Punishment points from subject j to i in period t-1		0.0484 (0.0075)***
α_i	-1.6996 (0.6983)**	-0.0217 (0.1091)
β_i	12.5428 (3.5425)***	0.0912 (0.6211)
Dummy for period $t = 10$	-2.5828 (0.5067)***	-0.2372 (0.0681)***
for $c = 0.5$		0.1502 (0.0580)**
for MIX	-0.5082 (0.4596)	-0.1333 (0.0673)**
for FAIR	0.7522 (1.2529)	0.2080 (0.8407)
for FAIR(ni)	1.7977 (1.2516)	-0.2382 (0.2328)
(Dummy for MIX) $\cdot \beta_i$	-9.9594 (3.5442)***	-0.2348 (0.6252)
(Dummy for FAIR) $\cdot \beta_i$	-12.2170 (3.9308)***	-0.1241 (1.3087)
(Dummy for FAIR(ni)) $\cdot \beta_i$	-15.7176 (4.1000)***	-0.2382 (0.2323)
	N = 1296 Prob > chi^2 = 0.0000 Pseudo R^2 = 0.2169	N = 1296 Prob > chi^2 = 0.0000 Pseudo R^2 = 0.2097

Notes: The coefficients are marginal effects on the unconditional expected value of the dependent variable. The marginal effects are obtained running the "dtobit" command in Stata after the corresponding Tobit estimation. Standard errors are in parenthesis. * Denotes significance at the 10%, ** at the 5% level and *** at the 1% level. Only the dummy for period $t = 10$ is indicated, no other dummy for a period is significant in both estimations. The corresponding reference is $t = 2$ in both games. The dependent variable is censored at a lower bound (0) and an upper bound (10) in game C and at a lower bound (-12) and an upper bound (0) in game D.

Regarding game D we have to bear in mind that the dependent variable "punishment points from subject i to j in period t" has a negative sign. We observe that the contribution of the co-player in the current period has a positive effect on the allocated punishment points by subject i, i.e. higher contributions of j lead to lower punishment from i to j. Remarkably, stronger punishment received from the co-player triggers stronger punishment imposed on the co-player, i.e. we observe a kind of escalation or "negative reciprocity" with respect to punishment behavior. Contrary to the estimation for game C, in game D, there are no effects of α_i and β_i on the dependent variable. Furthermore, higher costs of punishment lead to lower punishment, which is a quite intuitive result. Finally, as one may expect given the results in the previous section, subjects punish significantly stronger in treatment MIX than in treatment EGO.

> *Result 9: In the heterogeneous treatment MIX, subjects punish each other significantly stronger than in the homogeneous treatment EGO. Higher costs for*

punishment lead to lower punishment. There are no effects of α_i *and* β_i *on punishment behavior.*

5 Summary and discussion

The paper is aimed at finding out about the effects of heterogeneous individual other-regarding preferences on individual behavior in social dilemma situations as specified by PG games with and without the opportunity to punish co-players. In order to do so, we used two simple, non-strategic games to elicit the extent of our participants' other-regarding preferences. Equipped with this knowledge, we formed groups of subjects to play the PG games. The groups differed with respect to their composition. In particular, we formed groups of “fair”, i.e. highly inequity avers subject, groups of “egoistic” subjects that are very little inequity averse, and “mixed” groups consisting of both types of subjects.

Our results differ from the results described in comparable papers. Firstly, the weight of disadvantageous inequity, α_i, varies very little throughout our subject pool, and has a median of zero. Therefore, only a small share of our participants (12%) meet F&S' condition that aversion against advantageous inequity cannot be higher than aversion against disadvantageous inequity. Secondly, our results show that the specific composition of groups significantly influences the subjects' performance in the PG games: As long as subjects are informed about the type of their opponent, “fair” groups contribute more to the public good than “egoistic” or “mixed” groups. It turns out that explicit information is a key factor for this difference in behavior: As long as “fair” subjects are not informed on the fact that their co-players are “fair”, too, they act like “egoistic” subjects. Only the explicit information that they are playing a “fair” co-player significantly enhances their contributions. Uninformed “fair” players are not significantly more cooperative than “egoistic” ones.

The option to punish co-players enhances individual contributions. Net of punishment received and costs of punishment imposed, however, only groups of “fair” players stand to benefit from the punishment option. Particularly the informed “fair” players benefit from the punishing opportunity. This, again, is a strong hint that the composition of the groups in the different PG games as well as the state of information matter heavily.

6 Conclusion

Our results show that even if we understand F&S' model as purely payoff based, i.e. without any consideration of reciprocity or intentions, at least the weight of aversion against favorable inequity has some descriptive power. Subjects with a high value of β_i contribute significantly more to the public good than other subjects – under the condition that their co-player has a similar parameter value and that the subjects are mutually informed about that. However, given our results and those in Blanco et al. (2006) there are a lot of open questions with respect to within-subject tests of theories for other-regarding preferences.[22] We close mentioning two of these issues which may be the subject of future research. Firstly, if "preferences" are sensitive to the method of elicitation, i.e. whether preferences are elicited with a "strategic" or "non-strategic" technique, one may ask which are the "true" preferences. Secondly, one may argue that a model of individual behavior which is based purely on monetary payoffs, such as F&S, allows to derive individual preferences in simple distributional games. The question is then, which descriptive power these preferences have in a "strategic" environment such as a PG game. Our results show that F&S' model has descriptive power – at least for one parameter.

---

[22] For a variety of other points on consistency and stability of other-regarding preferences, see Brosig et al. (2007).

Appendix

This appendix includes the following instructions (translation from German):

1. Instructions for Games A and B
2. Information about the co-player in Games C and D
3. Instructions for Game C
4. Instructions for Game D

1. Instructions for Games A and B

Welcome to our laboratory experiment!

Please read the instructions carefully within the next 10 minutes. If you have any questions please raise your hand. In this experiment you can earn money depending on your decisions and the other subjects' decisions. Your decisions in the experiment will be anonymous. Only the experimenter will get to know your identity, but that information will be treated confidentially.

In the experiment you will have to make decisions in two games, Game A and Game B. The rules of the games as well as the determination of payoffs will be explained in the following.

Rules for Game A and Game B

There are two players in Game A and Game B: Player 1 and Player 2. Player 1 has to solve the following decision problem: From a list of two pairs of payoffs (I and II) for Player 1 and Player 2, he has to select one or the other. Player 2 has no choice, i.e. he has to accept the decision made by Player 1.

In the experiment you will make the decisions as Player 1 for Game A as well as for Game B, i.e. you will have to decide which of both pairs of payoffs (I or II) for Player 1 and Player 2 you select. Each game implies 22 decisions for Player 1.

Example: Player 1 has selected Pair I, which gives him x € and the other player y €.

	Pair I			Pair II		
No.	choice	Player 1	Player 2	choice	Player 1	Player 2
1	×	x €	y €		v €	w €

Please bear in mind: You have to choose between Pair I and II which are different in Game A and Game B.

Determination of payoffs

Your payoff from the experiment will be computed as follows: All subjects in this room will be matched into pairs of subjects using a random draw. After this, it will be determined (again, randomly) whether a pair of subjects plays Game A or Game B. After the selection of the relevant game, a random draw selects which number from the payoff list (between No. 1 and 22) will be relevant. Finally a random draw decides which decision – yours or that of your co-player – determines the payoffs.

If your choice is decisive, you will get the payoff for Player 1 and your co-player will get the payoff for Player 2. In the other case, if your co-player's choice is decisive, you will get the payoff your co-player has assigned to you. That means you will get the payoff for Player 2 and your co-player the payoff for Player 1. You will get the respective amount of money in cash at the end of the experiment.

Please bear in mind: According to these rules, each of your decisions in Game A and Game B can be relevant for your payoff as well as your co-player's payoff from the experiment.

If you have read through the instructions and do not have any questions, please answer the following control questions. You find the tables with the pairs of payoffs on the following pages.

1. Suppose Game A, No. 3, was selected. You have chosen Pair I and your co-player Pair II. Your choice determines the payoffs. What are the payoffs?

My payoff: _______ My co-player's payoff: _______

2. Suppose Game B, No. 13, was selected. You have chosen Pair II and your co-player Pair I. Your co-player's choice determines the payoffs. What are the payoffs?

My payoff: _______ My co-player's payoff: _______

3. Suppose Game B, No. 16, was selected. You have chosen Pair I and your co-player Pair I. Your choice determines the payoffs. What are the payoffs?

My payoff: _______ My co-player's payoff: _______

4. Suppose your decision determines the payoffs. Is it possible according to the applying payoff rule, that you get the payoff you have assigned to Player 2? Please mark your answer with a cross.

O yes, it is possible O no, it is not possible

Please give us a sign when your have answered all questions. We will then check your answers. The experiment will start as soon as all subjects have passed the test successfully. You will have to enter you decisions on the computer screen afterwards.

Good luck in the experiment!

The MaXLab team

Game A

In the experiment you have to mark your decision for Pair I or Pair II with a cross. Therefore, you have to decide 22 times which of both pairs of payoffs you choose for you and your co-player.

Your decision as Player 1

– payoffs in € –

	Pair I			Pair II		
No.	choice	Player 1	Player 2	choice	Player 1	Player 2
1		5.00	5.00		2.00	2.00
2		4.44	5.56		2.00	2.00
3		4.42	5.58		2.00	2.00
4		4.39	5.61		2.00	2.00
5		4.36	5.64		2.00	2.00
6		4.32	5.68		2.00	2.00
7		4.29	5.71		2.00	2.00
8		4.24	5.76		2.00	2.00
9		4.19	5.81		2.00	2.00
10		4.14	5.86		2.00	2.00
11		4.07	5.93		2.00	2.00
12		3.92	6.08		2.00	2.00
13		3.86	6.14		2.00	2.00
14		3.81	6.19		2.00	2.00
15		3.68	6.32		2.00	2.00
16		3.53	6.47		2.00	2.00
17		3.33	6.67		2.00	2.00
18		2.85	7.15		2.00	2.00
19		2.72	7.28		2.00	2.00
20		2.22	7.78		2.00	2.00
21		1.43	8.57		2.00	2.00
22		0.10	9.90		2.00	2.00

Game B

In the experiment you have to mark your decision for Pair I or Pair II with a cross. Therefore, you have to decide 22 times which of both pairs of payoffs you choose for you and your co-player.

Your decision as Player 1

– payoffs in € –

	Pair I			Pair II		
No.	choice	Player 1	Player 2	choice	Player 1	Player 2
1		10.00	0.00		0.00	0.00
2		10.00	0.00		0.50	0.50
3		10.00	0.00		1.00	1.00
4		10.00	0.00		1.50	1.50
5		10.00	0.00		2.00	2.00
6		10.00	0.00		2.50	2.50
7		10.00	0.00		3.00	3.00
8		10.00	0.00		3.50	3.50
9		10.00	0.00		4.00	4.00
10		10.00	0.00		4.50	4.50
11		10.00	0.00		5.00	5.00
12		10.00	0.00		5.50	5.50
13		10.00	0.00		6.00	6.00
14		10.00	0.00		6.50	6.50
15		10.00	0.00		7.00	7.00
16		10.00	0.00		7.50	7.50
17		10.00	0.00		8.00	8.00
18		10.00	0.00		8.50	8.50
19		10.00	0.00		9.00	9.00
20		10.00	0.00		9.50	9.50
21		10.00	0.00		10.00	10.00
22		10.00	0.00		10.50	10.50

2. Information about the co-player in Games C and D

Your booth: ….

Game A

Your co-player has always chosen Pair I from no. 1 to …. in Game A. He/She has always chosen Pair II from no. … to no. 22.

Game B

Your co-player has always chosen Pair I from no. 1 to …. in Game B. He/She has always chosen Pair II from no. … to no. 22.

3. Instructions for Game C

Rules of Game C

Game C consists of 10 separate rounds. In each round you will play the same game. The subject with whom you will interact will stay the same in each round. However, you will interact with a different subject than in Games A and B. Your co-player will not find out your identity and vice versa.

The payoff rule for Game C is the following: In the beginning of each round you dispose of 10 €. We denote this amount as endowment in the following. Your task (as well as your co-player's task) is to decide how much of your endowment you contribute to a project. The amount you don't contribute to the project, you keep for yourself.

In each round the payoff consists of two parts:

- The amount you keep and
- your income from the project.

The income from the project for a subject is calculated by multiplying the sum of the contributions of both subjects by 0.7, i.e.

Income from the project for a subject = 0.7 * (sum of contributions of both subjects)

For instance: If both subjects contribute 5 € to the project, the income from the project for both players is 7 € = (0.7 * (5 + 5)) for each. If one player contributes 10 € and the other one nothing, the income from the project is 7 € for both subjects, too. Every Euro you contribute to the project increases your income and your co-player's income by 0.70 € each. The same applies for the contribution made by your co-player. Every Euro he/she contributes increases her/his income from the project and your income from the project by 0.70 €.

The amount you keep for yourself equals your endowment minus your contribution to the project. The balance of 3 Euro will be added to your payoff in each round.

Example: Subject 1 contributes y € and Subject 2 contributes z € to the project.

	Subject 1	Subject 2
Endowment	10 €	10 €
Contribution to the project	y €	z €
Income from the project	0.7 * (y + z) €	0.7 * (y + z) €
Payoff	10 – y + 0.7 * (y + z) €	10 – z + 0.7 * (y + z) €
Payoff with balance	13 – y + 0.7 * (y + z) €	13 – z + 0.7 * (y + z) €

Information in Game C

At the end of each round, you will be informed about your contribution and you co-player's contribution. Additionally, you will find out your payoff and your co-player's payoff in each round.

Payoff in Game C

The payoff in Game C will be determined as follows:

After Game C, you will play game D. After Game D, a random draw will decide which game (C or D) will be is the relevant game for the payoff.

If Game C is selected, a random draw will select one of the 10 rounds. You will receive the payoff according to your decision and your co-player's decision in this round. Please bear in mind: Each round in Game C has the same probability to be selected as the relevant round.

Control questions for Game C

If you have read the instructions and don't have any questions, please answer the following control questions.

Suppose you and your co-player contribute each 0 € to the project. What are the payoffs with balance in this round?

My payoff: _______ My co-player's payoff: _______

Suppose you contribute 0 € to the project and your co-player contributes 5 €. What are the payoffs with balance in this round?

My payoff: _______ My co-player's payoff: _______

Suppose you and your co-player contribute each 10 € to the project. What are the payoffs with balance in this round?

My payoff: _______ My co-player's payoff: _______

Suppose you contribute 10 € to the project and your co-player contributes 0 €. What are the payoffs with balance in this round?

My payoff: _______ My co-player's payoff: _______

Once you have answered all questions, give us a sign. We will then check your answers. Game C will start as soon as all subjects have passed the test successfully.

Good luck in the experiment!

The MaXLab team

4. Instructions for Game D

Rules of Game D

Game D consists of 10 separate rounds. In each round you will play the same game. The subject with whom you will interact is the same as in Game C. As before, your co-player will not find out your identity and vice versa.

The payoff rule for Game D is the following: In the beginning of each round you dispose of 10 €. We denote this amount as endowment in the following. Game D has two stages. At stage 1, your task (as well as the task of your co-player) is to decide how much of your endowment you contribute to a project. The amount you don't contribute to the project, you keep for yourself.

Therefore, in each round the payoff at stage 1 consists of two parts:

- The amount you keep and
- your income from the project.

The income from the project for a subject is calculated by multiplying the sum of the contributions of both subjects by 0.7, i.e.

Income from the project for a subject = 0.7 * (sum of contributions of both subjects)

Therefore, stage 1 of Game D is equivalent to Game C, which you have played before. In each round of Game D, stage 1 is followed by stage 2. At this stage you have the possibility to assign your co-player negative points. A negative point reduces your co-player's payoff by 1 €. To assign negative points is costly for you. If you assign a negative point to your co-player, your payoff will be reduced by 0.50 €.

Therefore, your payoff with balance in each round is determined as follows:

- The payoff with the balance of 3 € at stage 1,
- minus the negative points you receive from your co-player
- minus the costs of the assignment of negative points by yourself.

Example: Subject 1 receives y € and Subject 2 z € from stage 1. Subject 1 assigns 6 negative points to Subject 2, Subject 2 assigns 0 negative points to Subject 1.

	Subject 1	Subject 2
Payoff with balance at stage 1	y €	z €
Received negative points	0 points	6 points
Assigned negative points	6 points	0 points
Costs of assignment of negative points	0.5 * 6 €	0 €
Payoff with balance	y – 3 €	z – 6 €

Information in Game D

You will be informed at the end of stage 1 about your contribution and the contribution made by your co-player. Additionally you will find out your payoff and your co-player's payoff at stage 1. You will be informed about the payoff reduction through assigned and received negative points for both subjects at the end of stage 2. Furthermore, the payoff at the end of stage 2 will be indicated.

Payoff in Game D

The payoff in Game D will be determined as follows:

After Game D, a random draw will decide which game (C or D) will be the relevant game for the payoff.

If Game D is selected, a random draw will select one of the 10 rounds. You will get the payoff according to your decision and your co-player's decision in this round. Please bear in mind: Each round in Game D has the same probability to be selected as the relevant round.

Control questions for Game D

If you have read the instructions and don't have any questions, please answer the following control questions.

Suppose you and your co-player contribute each 0 € to the project at stage 1. None of the subjects assigns negative points at stage 2. What are the payoffs with balance in this round?

My payoff: _______ My co-player's payoff: _______

Suppose you contribute 0 € to the project and your co-player contributes 10 €. Your co-player assigns 10 negative points to you at stage 2. You don't assign negative points. What are the payoffs with balance in this round?

My payoff: _______ My co-player's payoff: _______

Once you have answered all questions, give us a sign. We will then check your answers questions. Game D will start as soon as all subjects have passed the test successfully.

Good luck in the experiment!

The MaXLab team

References

Blanco, M., Engelmann, D., and Normann, H.-T. (2006): A within Subject Analysis of Other-Regarding Preferences, Royal Holloway College, University of London, working paper, October 4, 2006.

Bolton, G.E. and Ockenfels, A. (2000): ERC. A Theory of Equity, Reciprocity and Competition, American Economic Review, 90, 166-193.

Brosig, J., Riechmann, T., and Weimann, J. (2007): Selfish in the End? An Investigation of Consistency and Stability of Individual Behavior, FEMM Working Paper No. 07005, University of Magdeburg, Germany, 2007.

Camerer, C.F. (2003): Behavioral Game Theory. Experiments in Strategic Interaction, Princeton University Press, Princeton, New Jersey.

Carter, J. R. and Irons, M. D. (1991): Are Economists Different, and if so, why?, Journal of Economic Perspectives, 5, 171-177.

Dufwenberg, M. and Kirchsteiger, G. (2004): A Theory of Sequential Reciprocity, Games and Economic Behavior, 47, 268-298.

Falk, A. and Fischbacher, U. (2006): A Theory of Reciprocity, Games and Economic Behavior, 54, 293-315.

Fehr, E. and Gächter, S. (2000): Cooperation and Punishment in Public Goods Experiments, American Economic Review, 90, 980-994.

Fehr, E. and Gächter, S. (2002): Altruistic Punishment in Humans. Nature 415, 137-140.

Fehr, E. and Schmidt, K.M. (1999): A Theory of Fairness, Competition, and Cooperation, Quarterly Journal of Economics, 114, 817-868.

Fischbacher, U. (1999): z-Tree: A Toolbox for Readymade Economic Experiments, Working Paper No. 21, University of Zurich, 1999.

Forsythe, R., Horowitz, J.L., Savin, N.E., and Sefton, M. (1994): Fairness in Simple Bargaining Experiments, Games and Economic Behavior, 6, 347–369.

Frank, R.H., Gilovich, T.D., and Regan, D.T. (1993): Does Studying Economics Inhibit Cooperation?, Journal of Economic Perspectives, 7, 159-171.

Frank, R.H., Gilovich, T.D., and Regan, D.T. (1996): Do Economists Make Bad Citizens?, Journal of Economic Perspectives, 10, 187-192.

Güth, W., Schmittberger, R., and Schwarze, B. (1982): An Experimental Analysis of Ultimatum Bargaining, Journal of Economic Behavior and Organization, 3, 367-388.

Kahneman, D., Knetsch, J.L., and Thaler, R.H. (1986): Fairness and the Assumptions of Economics, Journal of Business, 59, 85–300.

Ledyard, J. O. (1995): Public Goods. A Survey of Experimental Research, in Kagel, J. H. and Roth, A. E. (Eds.), The Handbook of Experimental Economics, Princeton University Press, Princeton, NJ, 111-194.

Marwell, G. and Ames, R.E. (1981): Economists Free Ride, Does Anyone Else? Experiments on the provision of public goods IV, Journal of Public Economics, 15, 295-310.

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

Sefton, M., Shupp, R., and Walker, J. (2002): The Effect of Rewards and Sanctions in Provision of Public Goods, Discussion Paper, forthcoming in Economic Inquiry.