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

Astrid Dannenberg, Thomas Riechmann, Bodo Sturm, and Carsten Vogt

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

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

Astrid Dannenberg, Thomas Riechmann,
Bodo Sturm, and Carsten Vogt

Non-Technical Summary

There is a growing number of stylized facts which contradict the model of rational payoff maximizing actors in economics. Individual contributions to public goods such as climate protection projects are prominent examples for behavior not in line with standard economic theory. In those cases, people cooperate although it seems not rational to do so. The contradiction between standard economic models of selfish behavior and empirical observations has been a challenge for both theorists and experimentalists. In the last ten years a number of theories have been developed which try to close this gap in explanatory power. 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 disparity between standard game-theoretical predictions and experimental observations by altering the underlying utility function of subjects, but stick to the assumption that subjects behave rationally.

This study aims to investigate the additional explanatory power of models with other-regarding preferences. On the basis of a laboratory experiment we present a simple two-steps procedure for a within-subject test of the inequity aversion model of Fehr and Schmidt (1999). In the first step, subjects play selected games in order to estimate their individual other-regarding preferences and are, thereupon, classified according to their behavior. In the second step, subjects with specific preferences as defined by the Fehr and Schmidt model are matched into pairs and interact 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 pairs 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.

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), Mannheim

[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

July 2007

Abstract

We present a simple two-steps procedure for a within-subject test 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.

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 close this gap in explanatory power have 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 affects 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.

This study aims to investigate the additional explanatory power of models with other-regarding preferences. Thereby, 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 which is still solely based on monetary payoffs. 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. Furthermore, we control the information subjects receive about the type of their co-player.

Blanco et al. (2006) were the first who employed the two games that help to elicit the individual weights of F&S inequity aversion. Our approach is based on the method introduced by Blanco et al., but differs from their approach in some aspects (see section 3 for a discussion).

4

Our results differ from the results described in comparable papers in respect of two aspects. 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 co-player, "fair" groups contribute more to the public good than "egoistic" or "mixed" groups. Moreover, "fair" groups are more likely to cooperate in the final period of the public good games than other 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.

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 summarizes and discusses our results. Section 6 concludes and gives a brief outlook on further research.

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:

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

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

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

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 modified ultimatum and dictator games (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 (PG) 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 $k/(n-1) \geq a/2$, there exists a unique equilibrium with $g_i = 0\ \forall i \in \{1,...,n\}$.

3. If for all players $j \in \{1,...,n\}$ with $a + \beta_j > 1$ the condition

$$k/(n-1) < (a + \beta_j - 1)/(\alpha_j + \beta_j) \tag{3}$$

holds, 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. F&S show that the existence of a group of so called "conditionally cooperative enforcers" may enhance the prospects for cooperation. These individuals $i \in [1,...,n']$ 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 marginal costs of punishment. If

$$c < \frac{\alpha_i}{(n-1)(1+\alpha_i)-(n'-1)(\alpha_i+\beta_i)} \quad \forall i \in [1,...,n'] \tag{4}$$

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 $j \in [1,...,n']$ carries out some punishment on player i $p_{ji} = (g-g_i)/(n'-c)$ while all other players do not punish.

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 k, the number of players 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 the

---

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

case of two players, following condition (3), a "fair" player with preference $\beta_j > 1 - a$ will choose to contribute to the provision of the public good if the following condition is met:[2]

$$E(k) < \frac{a + \beta_j - 1}{\alpha_j + \beta_j} \tag{5}$$

where $E(k)$ denotes the expected value of k. Obviously, for $n = 2$, k can only take on the values one or zero, i.e. $E(k)$ is the probability a "fair" player attaches to the possibility that his co-player is an "egoistic" type. 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, (5) reduces to $\beta_j > 1 - a$.

If $E(k) > 0$, however, for increasing values of α_j it becomes more difficult to fulfil (5), i.e. to ensure that the condition is still met, players must have higher values of β_j.[3] 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, 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.

Condition (5) can also be used to illustrate the effect of information. If "fair" players are informed about the behaviour of their opponents in former modified ultimatum and dictator games this increases their confidence about their opponent's type. Particularly, in this case they should expect their opponent to be a "fair" type with higher probability if he or she behaved accordingly in the former games. Technically, this means because $E(k)$ decreases, condition (5) is more easily met for informed subjects. Hence, we should observe a higher

---

[2] If one skips F&S's assumption that $\alpha_i \geq \beta_i$, which might be appropriate for our subject pool (see section 4.1), then the additional condition $E(k) \leq a/2$ must hold. See Fehr and Schmidt (1999), p. 862, for details.

[3] This can be easily seen when (5) is solved for β_j. In this case, condition (5) reads $\beta_j > \frac{E(k)}{1 - E(k)} \alpha_j + \frac{1 - a}{1 - E(k)}$.

level of cooperation in groups consisting of informed "fair" players than in groups consisting of uninformed "fair" players.

3 Experimental design

3.1 Games

We used four different games (games A, B, C, and D) in our experimental design. Thereby, the only purpose of games A and B which were originally introduced by Blanco et al. (2006) was to measure each subject's preferences according to the F&S model. After these games, some of the subjects with certain preferences played games C and D in various treatments (see section 3.2). The design of the games is presented in the following.[4]

Game A is designed to measure the subjects' aversion against disadvantageous inequity. The game resembles the responder's basic decision situation in the ultimatum game but abstracts from strategic interaction, such that we can rule out individual behavior caused by strategic considerations such as intentions or reciprocity.[5] In this game, 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 himself or 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.[6] 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.

---

[4] See the appendix for the instructions that we distributed to our participants.
[5] The difference to the payoffs of the original ultimatum game is the fact that the conflict point payoffs (in €) are changed to (2, 2) instead of the original (0, 0).
[6] In the following we assume rational behavior of all subjects.

Regarding the first condition, a subject with aversion against disadvantageous inequity 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 he or she could attain an equal allocation with higher own payoff by choosing pair I. A subject who chooses pair II in #22 only or never switches 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).[7]

Game B – which resembles the decision problem in the dictator game – is designed to measure the subjects' aversion against advantageous inequity.[8] Again, each subject had to decide between two pairs of payoffs (pair I and pair II) each with an amount of money for himself or 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 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.

---

[7] 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 $-0.08 \leq \alpha_i \leq 0.04$.

[8] Strictly speaking, game B is equivalent to the dictator game only for decision #11. However, similar to the dictator game, game B creates a trade-off between own monetary payoff which creates advantageous inequity and a lower but equally distributed payoff.

Table 1: Payoffs in game A and game B

	#	game A pair I payoffs (in €) for player 1	2	pair II 1	2	α_i	game B pair I payoffs (in €) for player 1	2	pair II 1	2	β_i
	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
switching point from pair I to pair II	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

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 $0 \leq \beta_i < 1$. 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 arranged 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 $\beta_i \geq 1$, 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. Similar to game A, we can determine the upper and lower bounds for the individual's β_i by means of a subject's switching point. We

approximate the individual value of β_i by choosing the mean of the corresponding interval (see Table 1).[9]

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, among others, an ultimatum game to derive α_i, and a modified dictator game to derive β_i. While the modified dictator game they used is practically identical to game B in this study, our game A differs from the ultimatum game with the strategy method in Blanco et al. in an important aspect.[10] In the ultimatum game, each subject reacts to a specific proposal of his or her co-player, i.e. there is a distinct element of strategic interaction in this game. In our game A, however, each subject himself or 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. we use only "non-strategic games"[11] for the elicitation of both F&S parameters. 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 still may influence the behavior of subjects (see section 4.1).

Fehr and Schmidt (2005) recommend using "strategic games" in order to elicit preference parameters that capture not only traits of inequity aversion but, moreover, strategic considerations like intentions or reciprocity. As our main focus is on the effect of inequity aversion, we do not follow this recommendation. Our games are explicitly of non-strategic nature. One more remark on the methodology to elicit individual F&S preference parameters should be in order. A basic assumption underlying our design of the games A and B is that individuals are only driven by preferences of the F&S-type. This means in particular that individuals do not hold any specific preferences with regard to "efficiency". An example will help to illustrate this point. In our game B, the sum of both payoffs in pair II rises from €0 in #1 to €21 in #22. An individual caring for efficiency only will switch from pair I to pair II after #10 or after #11. Given an individual cares for equity and efficiency, we underestimate

---

[9] 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 $-0.05 \leq \beta_i \leq 0$.

[10] Game B in the present study differs in a small but important aspect of the corresponding game in the Blanco et al. study. In our game B, selfish subjects have to switch from pair I to pair II in the last decision row (#22). The same holds for game A (#21 and #22). This feature enables us to select subjects who do not behave consistently neither according to F&S nor to the standard theory of selfish behavior (as a special case of F&S).

[11] Following Fehr and Schmidt (2005, p. 47), we regard such games as "strategic" where each player has an influence on each other player's material payoff.

his or her aversion against advantageous inequity for $\beta_i > .53$ and overestimate it for $\beta_i < .53$.

Game C is a standard two-player PG game with a voluntary contribution mechanism.[12] 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 himself or herself raises the individual 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 is constant and equal to 0.7. The game was played using a partner design over 10 rounds with the number of rounds as common knowledge. After each round subjects were informed about their 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 the same two-player PG game as described above. Stage 1 in game D is followed by a stage 2. In this stage subjects have the possibility to assign his or her co-player negative points, i.e. a punishment mechanism (e.g. Fehr and Gächter 2000, 2002) 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 either €0.17 or €0.50 (depending on the treatment). Again, the game was played using a partner design over 10 rounds with the number of rounds as common knowledge.[13]

Subjects were paid separately for games A and B and games C and D. The payments from games 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 would be 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) would be relevant. Finally, a random draw decided which person within a pair determined the payoffs, i.e. whose decision as player 1 was realized. 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

---

[12] Blanco et al. (2006), too, played a standard two-player PG game (in a one-shot version) after eliciting the individual inequity aversion weights.

[13] 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 been set to zero). Luckily, though, a case like this did not occur during the experiments.

a quiz before the experiment started. Subjects who behaved consistently in games A and B (which had been played at first) were invited to play games C and D.[14] The payments from games C and D were determined in a similar way: After both games, a random draw determined which game (C or D) would be 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.

3.2 Subject pool and treatments

We ran 25 sessions with 18 to 20 subjects in each session. All in all, 492 subjects participated in games A and B and 160 of these subjects were invited to play games C and D. Sessions lasted about 90 minutes and the average earning was €5.90 for games A and B and €15.20 for games C and D. The sessions were conducted in November 2006 and May 2007 at the Magdeburg Experimental Laboratory (MaXLab). The experiment was fully computerized and anonymous.[15] In the laboratory, 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 (41%) is quite high. Almost two thirds of our subjects already had experience in experiments, i.e. have participated in at least one experimental session before.

Table 2: Subject pool – descriptive statistics

	absolute frequency	relative frequency in percent
total	492	100.0
consistent choices	371	75.4
study		
Management Science (MS)	173	35.2
Economics (Ec)	39	7.9
other Economics (othEc)	79	16.1
MS or Ec or othEc (Econ)	291	59.1
other than Econ (NonEcon)	201	40.9
experienced (at least one experiment before)	339	68.9
male subjects	270	54.9
female subjects	222	45.1

---

[14] 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 this game was to avoid a zero payment for subjects who didn't play games C and D (there was no show-up-fee). After this game these subjects were paid and left the laboratory. The mean payment (€3.60) was accordingly low. In the following we do not refer to the results of this game.

[15] We used Z-tree for programming. See Fischbacher (2007).

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 (for 40 of 492 subjects), we had subjects play game B first, then A. Moreover, in seven of our sessions (for 136 of 492 subjects), 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).[16] Therefore, we felt free to pool the data of all 25 sessions of games A and B conducted.

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.[17] Treatments differ with respect to (1.) the preference parameter β_i of the subjects that form a pair 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.

Table 3: Treatments in game C

Treatment	Parameter β_i, $i = 1, 2$	Information	Observations
EGO	$\beta_i < 0.3$	yes	35
MIX	$\beta_1 < 0.3 \wedge \beta_2 > 0.3$	yes	13
FAIR	$\beta_i > 0.3$	yes	17
FAIR(ni)	$\beta_i > 0.3$	no	15
Σ			80

Notes: (1.) All subjects in FAIR and FAIR(ni) have $\alpha_i = 0$. (2.) No one in our subject pool has $\beta_i = 0.3$. (3.) There is no difference in the distributions of β_i between FAIR and FAIR(ni) at the 5% level (K-S test).

Our four treatments in 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

---

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

[17] 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 (see section 4.1) we had to focus the definition of treatments on β_i.

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

All subjects who played game C also completed game D, the PG game with punishment possibilities. The different treatments in game D are displayed in Table 4. Within the EGO treatment, we distinguished between high and low costs of punishment. Therefore, pairs of subjects from the EGO treatment in game C were allocated into two separate EGO treatments in game D, EGO(h) and EGO(l). As in game C before, in most of the treatments, subjects were informed about their co-player's 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)	$\beta_1 < 0.3 \wedge \beta_2 > 0.3$	yes	high	10
FAIR(h)	$\beta_i > 0.3$	yes	high	17
FAIR(h, ni)	$\beta_i > 0.3$	no	high	15
Σ				77

Notes (see also Table 3): (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 $\beta_i > 0.3$ in MIX(h) fulfills condition (4).

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. In doing so we can exclude any repeated game effects, which may come into play if there is a repeated interaction between only two subjects. In the following we assume that, whenever F&S predict the existence of multiple equilibria, subjects play each of them with the same probability.

As there is no difference in the prognosis regarding the contribution behavior between games C and D we are able to formulate joint hypotheses for both games. Note that there is no treatment in game D where punishment is part of an equilibrium strategy: In MIX(h) none of the subjects with $\beta_i > 0.3$ fulfills condition (4), i.e. there are no subjects who may enforce cooperation. In FAIR(h) cooperation without punishment is an equilibrium because – in the

case of mutual cooperation – there are no incentives for subjects with $\beta_i > 0.3$ to punish each other. In this case, punishment would reduce the own payoff and create advantageous inequality.

Therefore, we can derive the following hypotheses regarding the contributions to the public good in different treatments for games C and D:

H1: In the EGO and MIX treatments of both games, zero contributions of all subjects should be observed.

H2: In the FAIR treatments, positive contributions to the public good should be observed more frequently than in the corresponding EGO and MIX treatments.

H3: In the FAIR treatments with information, positive contributions to the public good should be observed more frequently than in the corresponding FAIR(ni) treatment without information.

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.
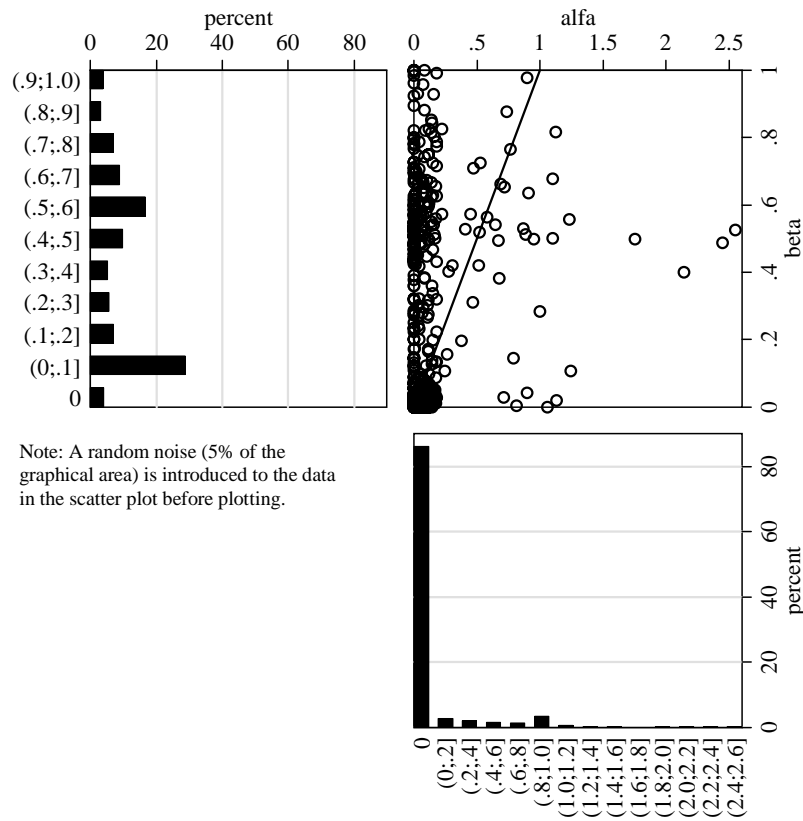

4.1 Behavior in games A and B

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 492 subjects who participated in the experiment 371 (75%) 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 about 86% of all subjects behave selfishly in game A (lower right in Figure 1). There is a second peak in the range of $0.8 < \alpha_i \leq 1.0$. 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 is 0.102. The median value is 0 indicating 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 is 0.356, the median is equal to 0.375. 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 (45 of 371, i.e. 12% of consistent choices) meeting the 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 $\rho = 0.345$, p = 0.299).

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 small but significant negative correlation between subjects who study Management Science, Economics or a related field (e.g. "business information systems") and the value for β_i (Spearman's $\rho = -0.154$, p = 0.003). This observation is in line with the well-known fact that economics students behave in general more selfishly than other students (e.g. Marwell and Ames 1981, Frank et al. 1996, and Carter and Irons 1991). However, this is the only significant correlation between α_i or β_i and the socio-economic characteristics at the 5% level.
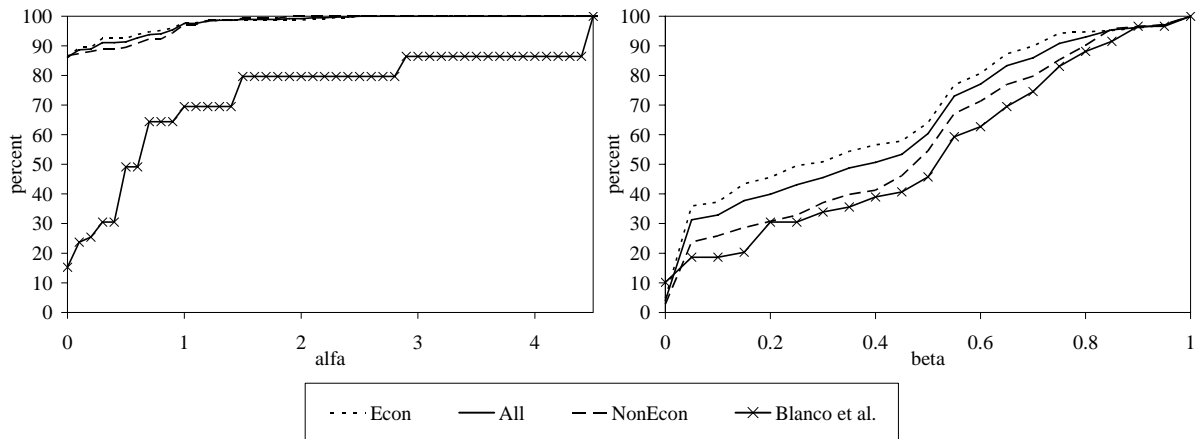
Figure 1: Distribution of F&S parameters for consistent choices

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

Because of the similarity between the approaches of revealing individual F&S preferences by Blanco et al. (2006) and our experimental technique, 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.[18] The difference in the cumulated densities for α_i between our subjects and the participants in Blanco et al. is remarkable and highly significant (K-S test, p = 0.000). While in Blanco et al. only 15% have $\alpha_i = 0$, in our study the fraction of subjects without any aversion against 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).

---

[18] We thank Hans Normann and his colleagues for providing the data of their experiment. All of the 72 subjects in Blanco et al. (2006) were non-economists. There were 13 subjects 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.

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

4.2 Behavior in games C and D

The results section for 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 estimates for the contribution to the public good in game C and the punishment behavior in game D.

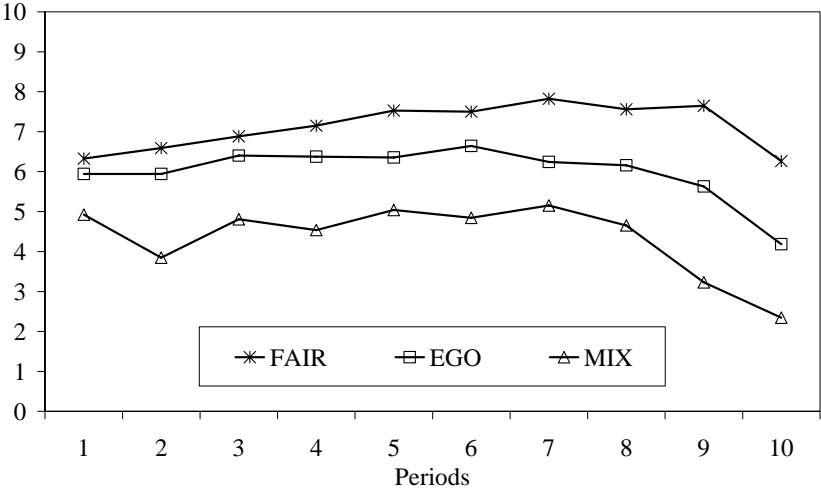### 4.2.1 Hypotheses testing of the F&S model in game C

A total of 160 subjects played the PG game C. The mean contribution per period over all subjects and all periods is €5.90. In order to test our hypotheses derived from the F&S model, we analyze the behavior of subjects in the treatments EGO, MIX, and FAIR (see Table 3). These treatments differ only in their composition of individual types, namely "egoistic", "fair" or both types. 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 solely depend on β_i.[19]

---

[19] If we control for α_i and apply the non-parametric tests to all subjects with $\alpha_i = 0$ (86% of our subjects in games C and D) we find the same results as in the tests including all subjects. Seeing this, α_i seems indeed to have no significant effect on subjects' contribution behavior. We will return to this point later.

The prognosis of the F&S model is the same for EGO and MIX. Contributions of zero are the only equilibrium. This is the same allocation that standard economic theory predicts for both games. In contrast, according to F&S in the FAIR treatment exist the "standard" non-contribution equilibrium as well as equilibria with positive contributions. Figure 3 shows the mean contributions in game C for the three treatments with information, namely FAIR, EGO, and MIX. The mean contribution per period is €7.10 for FAIR, €6.00 for EGO, and €4.30 for MIX. A Mann-Whitney U test (MW U test) shows that the differences between treatments are not significant at the 5% level except for FAIR and MIX (p = 0.013). 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). Therefore, we can sum up these observations into the following result.

> *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.*

Figure 3: Contributions in game C

As already explained above (section 3.3), our analysis focuses particularly on the behavior in the final period. The mean contribution in the final period is €6.30 in FAIR, €4.20 in EGO, and €2.40 in MIX. Thereby, the differences in final period contributions are significant between MIX and FAIR (MW U test, p = 0.007) and weakly significant between EGO and FAIR (p = 0.090). There are, in contrast, no significant differences between EGO and MIX. In consideration of our hypothesis *H2*, this is what we expected. According to this, the F&S

model seems to have 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 consider the share of subjects who "defect" in the final period by contributing nothing to the public good. 54% of the subjects in EGO, 58% of the MIX subjects but only 32% of the FAIR subjects defect in the final period. Thus, the shares of defecting subjects are relatively high in EGO and MIX compared to the FAIR treatment. Applying a χ^2 test, we find that these differences are not significant at the 5% level. One has to take into consideration, however, that the χ^2 test refers to independent observations, i.e. mean contribution of both players in a pair, and not to individuals. As a consequence, defection is only counted if both players in a pair contribute nothing. If one of the two players contributes a small amount to the public good, mean contribution is positive and, therefore, represents cooperation in the test. For this reason, we repeat the test with the critical value for defection equal to €3. Thus, we define defection when the mean contribution of both players in a pair is below €3 and cooperation when it is equal or above €3. In this case, the differences in shares of defecting subjects is significant between MIX and FAIR (p = 0.018) and weakly significant between EGO and FAIR (p = 0.063). Again, there are no significant differences between EGO and MIX. Furthermore, we can reject the hypothesis that cooperation and defection are equiprobable for FAIR (p = 0.008) whereas we cannot reject this hypothesis for EGO and MIX at the 5% level. This holds regardless of whether we take zero contributions or contributions below €3 as defection.

> *Result 2: The F&S model has some 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 EGO and MIX. Furthermore, the share of subjects who "defect" in the final period is lower in FAIR than in EGO and MIX.*
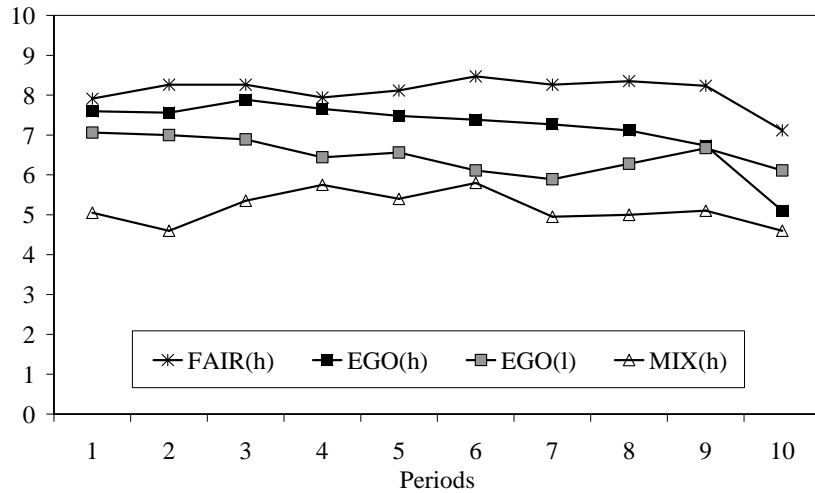
Figure 3 illustrates the relatively poor performance of the subjects in the MIX treatment. Even though the differences between MIX and EGO are not significant, it is remarkable that the contributions in MIX are always lower. The MIX treatment is the only heterogeneous treatment where two different types 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 MIX contribute on average €6.30 in the first period whereas the "fair" subjects contribute only

€3.50. 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, which prevents the "fair" subjects from being cooperative in the PG game. We will return to this aspect later.

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

All of the 160 subjects who played game C also completed game D, the PG game with punishment possibility. The mean contribution per period over all subjects and all periods is €6.90. In order to test our hypotheses for game D we consider the behavior of subjects in the treatments FAIR(h), EGO(h), and MIX(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 EGO(h) and MIX(h) whereas equilibria with positive contributions are also feasible for FAIR(h). Figure 4 shows the development of the mean per-period contributions over time for all three treatments. The mean contribution over all periods for FAIR(h) is €8.10, for EGO(h) €7.20, and for MIX(h) €5.20. The differences between treatments are not significant at the 5% level except for those between FAIR(h) and MIX(h) (MW U test, p = 0.013). In all periods and in all treatments, the share of cooperating subjects who contribute a positive amount to the public good is significantly greater than zero (Binomial Sign test, one-tailed, p = 0.000), indicating that the contributions of the subjects in EGO(h) and MIX(h) 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.
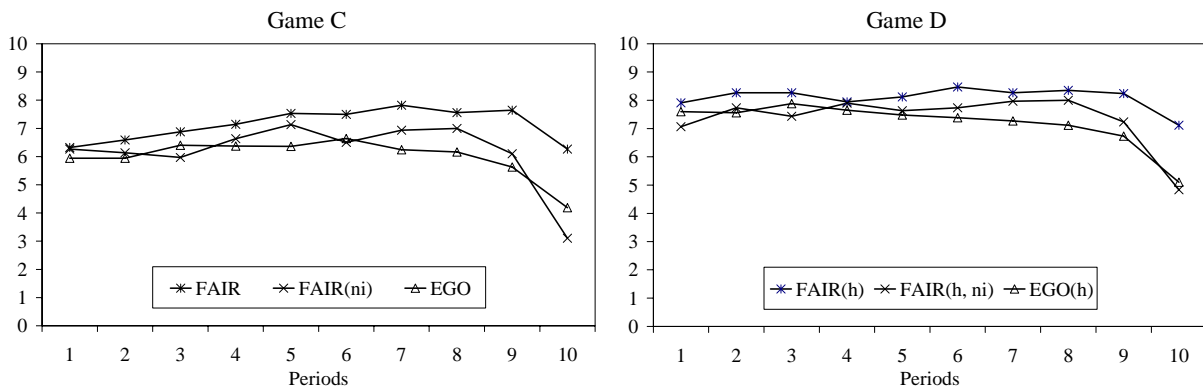
Again, we consider the subjects' behavior in the final period in more detail. The mean final period contribution in game D is €7.10 for FAIR(h), €5.10 for EGO(h), and €4.60 for MIX(h). This time, however, the differences between treatments are insignificant at the 5% level (MW U test). Considering the share of defecting subjects, we find that 48% of all subjects in EGO(h) and 45% of the MIX(h) subjects contribute nothing in the final period compared to only 24% of the FAIR(h) subjects. As already observed in game C, pairs consisting of two "fair" players seem to be more likely to cooperate in the final period than "egoistic" or "mixed" pairs. The comparison between treatments via a χ^2 test exhibits that the share of defecting subjects does not significantly differ between EGO(h) and MIX(h) whereas it significantly differs between EGO(h) and FAIR(h) (p = 0.026) as well as between MIX(h) and FAIR(h) (p = 0.018). We can, furthermore, reject the hypothesis that defection and cooperation are equiprobable for FAIR(h) (p = 0.002) whereas we cannot reject this hypothesis for EGO(h) and MIX(h) at the 5% level. These results remain virtually the same when we take contributions below €3 instead of zero contributions as defection.

Result 4: The F&S model has some explanatory 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 share of defecting subjects is significantly lower in FAIR(h) than in the other treatments.

4.2.3 Effect of information

As mentioned above, except for the treatment FAIR(h, ni), subjects were informed about how their co-player behaved in games A and B previous to games C and D. As 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.[20]

Figure 5: Contributions – effect of information

The mean contributions of FAIR(h) and FAIR(h, ni) are shown in Figure 5. For reference, we also insert the contributions of the "egoistic" subjects in EGO and EGO(h). In both games and in all periods the contributions of the uninformed subjects are lower than the contributions of the informed subjects, though the differences between FAIR(h) and FAIR(h, ni) in mean per-period contribution are not significant at the 5% level (MW U test). In the final period of both games, however, the contributions in FAIR(h) and FAIR(h, ni) differ significantly (MW U test, game C: p = 0.025, game D: p = 0.048). Remarkably, the information about the co-player seems to play an important role mainly in the final period, when strategic incentives to

---

[20] All 64 subjects in FAIR and FAIR(ni) have $\alpha_i = 0$ and $\beta_i > 0.3$. There is no difference in the distributions of β_i between FAIR and FAIR(ni) at the 5% level (K-S test) (see Table 3).

cooperate do not exist any longer. In contrast, we do not find significant differences in final period contributions between FAIR(ni) and EGO as well as between FAIR(h, ni) and EGO(h). It seems that in the final period of both games, the uninformed "fair" subjects behave similar to the "egoistic" subjects rather than to the informed "fair" subjects. Furthermore, the contributions of the informed subjects and uninformed subjects do not converge over time (see Figure 5). Hence, even though the uninformed "fair" subjects have the chance to get to know each other through interaction in the PG games this does not lead to the same level of contributions as observed for informed subjects. Therefore, we can deduce that the information subjects receive about the type of their co-player before games C and D (based on the decisions in games A and B) is not equivalent to the information subjects are able to gather via direct interaction with their co-player during these PG games.

Let us now consider the share of defecting subjects in the final periods of both games. In game C, 67% of the FAIR(ni) subjects contribute nothing to the public good compared to only 32% of the FAIR subjects. Thus, the share of defecting subjects in FAIR(ni) is even larger than the share in EGO where 54% defect in the final period. A χ^2 test indicates that the difference in shares of defecting subjects between FAIR(ni) and EGO is not significant at the 5% level whereas it is highly significant between FAIR(ni) and FAIR (p = 0.004). This result applies to both critical values for defection, zero contributions and contributions below €3. In the final period of game D, 43% of the FAIR(h, ni) subjects contribute nothing to the public good compared to 24% of subjects in FAIR(h). The difference in defection between informed and uninformed subjects, however, is not significant anymore at the 5% level. This also holds when we take contributions below €3 as defection. To sum up, our observations regarding the behavior of the informed and uninformed subjects generally support hypothesis *H3*, although the evidence for an information effect is weaker in the game with punishment possibilities.

> *Result 5: In the final period of both games, uninformed "fair" subjects contribute significantly less than informed "fair" subjects. Furthermore, in the final period of the standard PG game the uniformed subjects are more likely to defect than the informed subjects.*

4.2.4 Effect of punishment

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 (Fehr and Gächter 2000, 2002) show, however, that the introduction of punishment usually increases the contributions. This is also what we observe in game D. The mean per-period contribution over all periods is €5.90 in game C and €6.90 in game D.[21] 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. 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 significant influence on the level of contributions to the public good.[22]

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 on 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. 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.[23]

---

[21] At this point, we forbear from a separate graphical presentation. However, the magnitude of the contribution enhancing effect of punishment is also visible in Figure 6, if one compares the efficiency in game C and the gross efficiency in game D.

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

[23] Note that net efficiency can even be negative. 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.
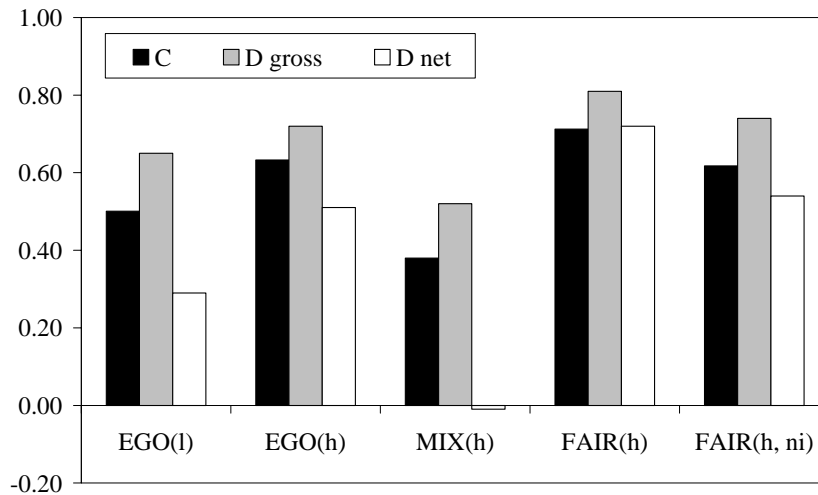
Figure 6: Efficiency

Figure 6 shows the mean degree of efficiency over all periods in games C and D for all treatments. As expected, gross efficiency in game D always exceeds efficiency in game C. Net efficiency, however, is lower than efficiency in game C for treatments EGO(l), EGO(h), MIX(h), and FAIR(h, ni). This is particularly compelling in MIX(h) where net efficiency is negative. The difference between gains from contributions only and these gains net of costs induced by punishment 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 only "fair" or "egoistic" types, 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 in contributions. As a consequence, we often witness strong escalation of punishment over time without any positive effects on contributions.[24]

The comparison between net efficiency of the treatments shows that differences are significant between MIX(h) and EGO(h) (MW U test, p = 0.017) as well as between MIX(h) and FAIR(h) (p = 0.003). Net efficiency between EGO(h) and FAIR(h) does not significantly differ at the 5% level. In contrast to all other treatments, the net efficiency of FAIR(h) in the punishment condition approximates the efficiency in the no-punishment condition. In this case, punishment does not lead to efficiency losses. Furthermore, the positive effect of information is still at work. The comparison between net efficiency shows that the difference between FAIR(h) and FAIR(h, ni) is significant (MW U test, p = 0.043) whereas it is not significant between EGO(h) and FAIR(h, ni). Thus, the behavior of the uninformed "fair"

---

[24] See also the Tobit estimates in the next section for this aspect.

subjects resembles the behavior of the informed "egoistic" subjects rather than the behavior of the informed "fair" subjects.

> *Result 6: The introduction of an option to punish the co-player significantly increases the mean contribution to the public good. Punishment decreases net efficiency compared to a no-punishment environment for all treatments except FAIR(h). The decrease is particularly distinctive in the heterogeneous MIX(h) treatment. The information effect is still at work.*

4.2.5 Multivariate analysis

In the sections before, we used 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 dependent 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 in the previous period has – over all treatments – a strong and significantly positive effect on the contribution of subject i in the current period. Parameter α_i has a negative effect on the contributions of subject i over all treatments which is weakly significant. This effect is quite surprising given the low dispersion for α_i in the data (see Figure 1). Furthermore, β_i has a positive effect on i's contribution, which is significant at the 5% level. Both findings are in line with F&S' model if one assumes that there is, over all treatments, some uncertainty about the co-player's type (see section 2.2.3). As one could expect from Figure 3, contributions in treatments MIX and FAIR(ni) are significantly below the contributions in the reference treatment (EGO).

> *Result 7: While α_i has a weakly significantly negative effect on the contributions to the public good over all treatments, the effect of β_i is significantly positive.*

Table 5: Tobit estimates

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	-2.1525 (0.4091)***	0.3044 (0.1158)***
Contribution of co-player j in period t		0.0980 (0.0108)***
Contribution of co-player j in period t-1	0.9560 (0.0403)***	
Contribution of subject i in period t		-0.0398 (0.0093)***
Punishment points from subject j to i in period t-1		0.0739 (0.0109)***
α_i	-1.0937 (0.6131)*	-0.0611 (0.1512)
β_i	1.7158 (0.7702)**	-0.1352 (0.1995)
Dummy for $c = 0.5$		0.1782 (0.0857)**
MIX	-0.9537 (0.4114)**	-0.2404 (0.0942)**
FAIR	-0.8372 (0.5782)	-0.1589 (0.1450)
FAIR(ni)	-1.5298 (0.5648)***	-0.0841 (0.1465)
	N = 1440	N = 1440
	Prob > chi^2 = 0.0000	Prob > chi^2 = 0.0000
	Pseudo R^2 = 0.2168	Pseudo R^2 = 0.1683

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. Dummies for periods are included but not indicated. 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 in the previous period 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. This result is in line with the F&S model: Given our subject pool and the composition of groups, there should be no effect of both parameters on the punishment behavior (see sections 2.2 and 3.3). 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 (see also Figure 5), subjects punish significantly stronger in MIX than in EGO.

Result 8: 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 significant effects of α_i and β_i on punishment behavior.

5 Summary and discussion

The present paper investigates the effects of heterogeneous 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 according to the F&S model. 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 averse 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 in respect of two aspects. 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%) meets the F&S condition that aversion against advantageous inequity cannot be higher than aversion against disadvantageous inequity. At this point, the difference in the α_i distribution compared to the above mentioned Blanco et al. study, which is (presumably) caused by the non-strategic design of our modified ultimatum game, is remarkable. Secondly, our results show that the specific composition of groups significantly influences the subjects' performance in the final period of the PG games: In the standard PG game, "fair" groups contribute more to the public good than "egoistic" or "mixed" groups. In both PG games (with and without punishing possibility) purely "fair" groups are more likely to cooperate in the final period than the other groups. In addition, it turns out that explicit information is a key factor for the difference in behavior: Uninformed "fair" subjects contribute in both PG games significantly less than informed "fair" subjects. In the final period of the standard PG game, the uniformed subjects are more likely to defect than the informed subjects. Thus, as long as "fair" subjects are not informed on the fact that their co-player is "fair", too, they act like "egoistic" subjects. The absence of any convergence of the behavior of informed and uninformed "fair" subjects

suggests that the information about the co-player's type cannot be extracted during the PG game. Only the explicit information that they are playing a "fair" co-player, which is given before the PG game, significantly enhances their contributions. Thus, we may conclude that the F&S model provides – at least as far the aversion against advantageous inequity is concerned – some additional explanatory power compared to the standard model of pure selfishness.

The option to punish co-players enhances individual contributions. This observation is consistent with the previous literature. However, our design allows us to investigate to what extend the introduction of a punishment possibility generates different effects on the payoffs of different types of subjects. Remarkably, net of punishment received and costs of punishment imposed, only groups of informed "fair" players are able to maintain their payoff from the no-punishment condition. All other groups incur payoff losses due to punishing. This, again, is a strong hint that the composition of the groups in the different PG games as well as the state of information about the co-player's type matter heavily.

One more remark on the potential policy relevance of the results presented in the paper is in order. Our findings can be particularly relevant for solutions of social dilemma problems, i.e. for situations where individually rational behavior does not lead to a socially optimal outcome. The behavior of subjects in our experiment indicates that the existence of fair types of subjects, i.e. who are reluctant to free ride at the expense of the others in the group, is only a necessary but not a sufficient condition for the solution of the social dilemma. Additionally, these fair subjects have to be informed by a credible institution about the fact that subjects in their group are of the same type. Without this ex ante information fair subjects are not able to solve the dilemma better than egoistic subjects.

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 advantageous inequity has some explanatory power for the individual behavior in social dilemma situations. In the standard PG game, 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. Given these conditions, subjects with a high β_i are also more likely to cooperate in both PG

games (with and without punishing possibility). 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.[25] 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 preferences are the "true" ones. Secondly, if one accepts that a model of individual behavior purely based on monetary payoffs, such as F&S, effectively allows deriving individual preferences with simple distributional games, the question is which explanatory power these preferences have in a "strategic" environment such as a PG game. Our results show that the F&S model has explanatory power – at least for one parameter.

---

[25] 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.

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

No.		Pair I			Pair II	
	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 €–

No.	Pair I choice	Pair I Player 1	Pair I Player 2	Pair II choice	Pair II Player 1	Pair II 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

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 will be added to your payoff in each round.

	Subject 1 (in €)	Subject 2 (in €)
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.

Fehr, E. and Schmidt, K.M. (2005): The Economics of Fairness, Reciprocity and Altruism - Experimental Evidence and New Theories, Discussion paper 2005-20, Department of Economics, University of Munich.

Fischbacher, U. (2007): z-Tree: Zurich Toolbox for Ready-made Economic Experiments, forthcoming in Experimental Economics, 171-178.

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.

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.