CESifo, the international platform of the University of Munich's Center for Economic Studies and the Ifo Institute for Economic Research

Venice Summer Institute 2005

Workshop on "Economics and Psychology"

18 – 19 July 2005

Venice International University, San Servolo

VENICE INTERNATIONAL UNIVERSITY

Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications

Simon Gächter

CESifo
Poschingerstr. 5, 81679 Munich, Germany
Phone: +49 (89) 9224-1410 - Fax: +49 (89) 9224-1409
office@CESifo.de
http://www.cesifo.de

Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications

Simon Gächter

University of Nottingham*, CESifo, IZA

July 2005 – VERY PRELIMINARY

1. The problem of voluntary cooperation

A well-known fact from the theory of public goods is that its voluntary provision will lead to an inefficient undersupply (Samuelson 1954). The reason is the well-known free rider problem: Since, by definition of a public good, an agent can benefit from it, even if he or she has not contributed to the public good, everyone has an incentive to hope that others provide the public good. More specifically, a rational and selfish agent will only equate his or her private marginal benefits and costs of the public good, whereas efficiency requires that the sum of marginal benefits should equal the marginal costs. Thus, there exists a tension between individual and collective rationality, which is prototypical for many cooperation problems. This tension lies at the heart of the matter in such diverse areas like warfare, environmental protection, management of commons, tax compliance, corruption, voting, the participation in collective actions like demonstrations and strikes, donations to charities, teamwork, collusion between firms, embargos and consumer boycotts, and so on.

While the logic of self-interest is straightforward, the facts seem to be at odds with the free rider hypothesis that is derived under the joint assumptions of rationality and selfishness. The fact that people even in anonymous situations vote, take part in collective actions, often manage not to overuse common resources, care for the environment, mostly don't evade taxes on a large scale, donate to public radio, as well as to charities, etc. *suggests* that the strict self-interest hypothesis is inconsistent with the degree of voluntary cooperation that we observe around us.

How can we explain this? What are the implications for public policy and management? This paper outlines some possible answers to both these questions. Yet, it is surprisingly hard

* School of Economics, Sir Clive Granger Building, University Park, Nottingham NG7 2RD. email: simon.gaechter@nottingham.ac.uk.

in the first place to rigorously test the free rider hypothesis in the field, since there are many factors operative at the same time, which preclude an *unambiguous* test of the free rider hypothesis. Therefore the lab is an important source of information to test whether, ceteris paribus, there is indeed more cooperation than predicted, when all other explanatory factors are held constant (see Ledyard 1995).[1]

The main finding from a large body of experiments that have been conducted in a variety of settings in the last three decades is that there is much more cooperation than predicted by standard theory (Ledyard 1995). Yet, the experiments also show that in repeatedly played public goods games, cooperation declines over time. In the next section I will substantiate these claims by presenting evidence from systematic experimental investigations – both from the lab and the field.

How can we explain (the decline of) voluntary cooperation? One important motive that has been discussed in the literature is that people have "warm-glow" preferences, i.e., they have some positive utility simply from the act of contributing (e.g., Andreoni 1990). A second explanation is that many people have altruistic preferences – they want to benefit others. A third reason is errors – people make mistakes (e.g., Anderson, Goeree and Holt 1998). In an ingenious design Palfrey and Prisbrey (1997) test for warm-glow, altruism and errors and find that altruism does not explain contributions, but some people have "warm-glow" preferences. Errors are important as well and explain why in repeated experiments contribution rates typically decline.

It should be noted that both motives – altruism and "warm glow" – are independent from other people's cooperation behavior. A set of recent experiments has cast doubt on this assumption. As I will show in the next section, a large number of people are "conditionally cooperative" – they cooperate if they believe others cooperate as well. Yet, a significant fraction of people is best characterized as free riders. In summary, recent evidence suggests that there is considerable heterogeneity with respect to people's cooperation preferences, i.e., there are "types" of players.

In section 3 I will take up the issue of preference heterogeneity and discuss four of its predicted consequences:

(i) Voluntary cooperation is fragile.

(ii) There exist social interaction effects in voluntary cooperation.

[1] The laboratory allows for a degree of control that is often not feasible in a naturally-occurring field situation. In all the experiments that we will discuss below participants, depending on their decisions, earned considerable amounts of money. Thus, the laboratory allows observing real economic behavior under controlled circumstances (see Kagel and Roth 1995 for an overview of experiments in economics.

(iii) Group composition with respect to "types" matters for voluntary cooperation.

(iv) Belief management matters for voluntary cooperation.

I will present evidence from new experiments that were set up to test these predictions.

First, without further institutional remedies, like punishment or assortative interactions, voluntary cooperation is predicted to be highly fragile. The reason is that conditional cooperators, who experience free riding, will stop cooperating themselves. Therefore, voluntary cooperation is bound to unravel.

Second, conditional cooperators will adapt their behavior to the respective group they are in. If other group members shirk, they shirk as well; if others cooperate people cooperate as well. In other words, there exists "social interaction effects", whereby people's behavior is influenced by their group mates.

Third, group composition with respect to "types" matters. For instance, if conditional cooperators know that the other group members are cooperators as well, then they should be able to maintain high cooperation levels. The "team spirit" of "like-minded" cooperators should suffice to maintain high cooperation. Similarly, free riders, who know that others are "free rider types" as well, are predicted to defect.

Fourth, by definition, conditional cooperators cooperate if they believe others cooperate as well. Hence, any factor that influences beliefs will affect cooperation behavior.

The evidence from these experiments unequivocally supports the importance of conditional cooperation and preference heterogeneity to understand cooperation behavior.

I see the experiments as behavioral models that may help us understand important field phenomena. In section 4 I will therefore quickly interpret field evidence on tax evasion, bribery, and welfare fraud, attitudes toward the welfare state, charitable giving, and work morale in the light of the four behavioral models.

The findings on the importance of conditional cooperation and preference heterogeneity have consequences for theory and policy. Take the theoretical reasons first. If people are largely motivated by 'warm-glow' preferences and if the decay in contributions is due to reduced errors, then the modeling approach might be another one than if people were free riders or conditional cooperators whose interaction explains the decay in contributions. In the former case, a modeling approach where errors figure prominently might be the preferable one (see, e.g., Anderson et al. 1998). In the latter case, a theory of social preferences might be chosen (see, e.g., Camerer 2003, and Fehr and Schmidt 2003 for surveys of models, and Tyran and Sausgruber for a policy application). The findings also have consequences for public policy and management. I discuss them in section 5. Section 6 concludes.

2. Conditional cooperation in the lab and the field

I start by presenting some stylized facts from laboratory experiments (section 2.1). This will only be a sketch and the interested reader may wish to consult Ledyard (1995) for a more complete account of important results from economic experiments. Dawes (1980) discusses evidence from social psychological experiments. In section 2.2 I will discuss a couple of recent field experiments that are consistent with the lab findings. Section 2.3 presents evidence that behavior in the lab is consistent with naturally occurring field behavior.

2.1 *Evidence from the laboratory*

An economic model of public goods provision that has proved extremely useful for testing the free rider hypothesis in the lab is the linear public goods game (or voluntary contribution mechanism). In a typical linear public goods experiment, n people form a group. All group members are endowed with z 'tokens'. Each subject i has to decide independently how many tokens (between 0 and z) to contribute to a common project (the public good). The contributions of the whole group are summed up. The experimenter then multiplies the sum of contributions by Ŭ and distributes the resulting amount equally among the four group members. Thus each subject i's payoff is

$$\phi_i \mid z 4 g_i 2 \frac{\zeta}{n} \underset{j \mid 1}{\overset{n}{\text{——}}} g_j, j \mid 1,...,n, \zeta \} 1, \zeta / n \{ 1. \qquad (1)$$

The first term ($z - g_i$) indicates the payoff from the tokens not contributed to the public good (the "private payoff"). The second term is the payoff from the public good. Each token contributed to the public good becomes worth Ŭ > 1 tokens. The resulting amount is distributed equally among the n group members – irrespective how much an individual has contributed. Thus, an individual benefits from the contributions of other group members, even if he or she has contributed nothing to the public good. Therefore, a rational and selfish individual has an incentive to keep all tokens for him- or herself, since the "return" per token from the public good for him- or herself is only Ŭ/n < 1, whereas it is 1 if he or she keeps the token. By contrast, since Ŭ > 1, the group as a whole is best off if everybody contributes all z tokens.

Figure 1 depicts a typical finding of a public goods experiment, where the exact same game is repeated ten times and subjects, who play in groups of four – know this. In each period subjects receive 20 tokens and decide how many of them to keep or contribute to the public good. After each round subjects are informed about what the other three group members have contributed. Figure 1 shows the resulting cooperation patterns in a "Stranger"

condition, where group members change randomly from round to round, and a "Partner" condition, in which groups stay constant for all rounds.

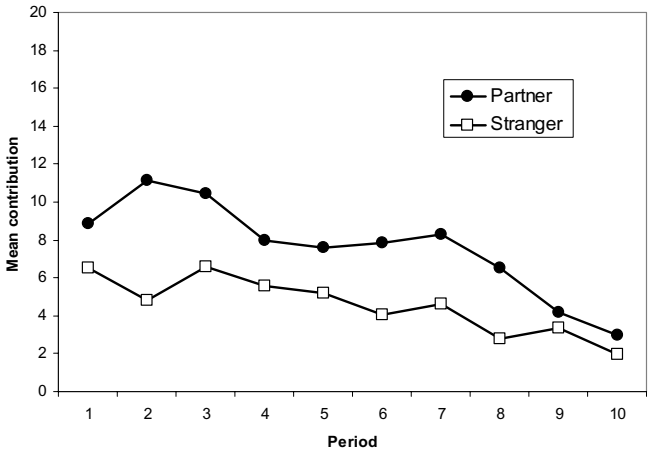

Figure 1: Contributions to a public good in constant ("Partner") and randomly changing groups ("Strangers") over ten repetitions. Source: Fehr and Gächter (2000).

Figure 1 illustrates two stylized facts from dozens of public goods experiments. First, people contribute substantially more than theoretically predicted. "Partners" contribute more than "Strangers" (see Keser and van Winden 2000, and Andreoni and Croson 1998 for an overview). The significance of this and related findings is that people are immediately able to distinguish whether they are in a situation that requires strategic cooperation (the "Partner" condition) or not (the "Stranger" condition) and to adapt their behavior accordingly.

A second stylized fact is that cooperation is very fragile and tends to collapse with repeated interactions. Why is this so? One explanation is that people have altruistic or "warm-glow" preferences, but also have to learn how to play this game. Since errors can only go in one direction, any erroneous decision looks like a contribution. Palfrey and Prisbrey (1997) test these explanations and find that the data are inconsistent with altruism. They find some evidence for "warm glow" but also conclude that over time, people learn and commit fewer errors, which is why contributions decline. The problem with this explanation is that it is inconsistent with the fact that after a so-called "restart" (after the tenth round participants are told that they will play another ten rounds) cooperation jumps up again and basically starts at the same level as in the first period. If learning would explain the decay in cooperation, then, after the restart, cooperation should have continued at the level at which cooperation was in the tenth round (see Andreoni 1988).

Notice that "warm glow", altruism and errors are motivations that are independent of others' contributions. Experiments by Croson (1998) were among the first to cast doubt on this. She simply elicited beliefs about other group member's contributions and found a very

high and statistically significant correlation of beliefs and contributions: Subjects who expected others to contribute a lot were more likely to contribute high amounts than subjects who expected others to free ride. This observation clearly suggests that people's contribution behavior is *not* independent of what they expect others to do. Thus, Croson's findings are evidence for "conditional cooperation".

Croson (2002) did not look at individual behavior. Her observation is that on average people behave conditionally cooperatively in that their contributions and beliefs are positively correlated. Fischbacher and Gächter (2005) also elicited beliefs and replicated Croson's findings of a positive correlation between beliefs and contributions. At the individual level they find subjects who do show a positive correlation between beliefs and contributions, whereas other subjects contribute zero even if they believe that others contribute positive amounts.

Fischbacher, Gächter and Fehr (2001) and Fischbacher and Gächter (2005) use a revealed preference method to infer people's contribution preferences in a public goods game as a function of other group members' contributions. Therefore, the subjects in their experiment do not choose one contribution but a contribution as a *function* of other group members' average contribution. The public goods game is played in groups of four subjects and the payoff function is again the same as in (1). The game is played just once to avoid confounds with strategic considerations. Every subject has to indicate a contribution *conditional on the average others' contribution*, i.e. for each of the 21 possible values of the average others' contribution subjects have to enter the number of points they want to contribute.

Fischbacher et al. (2001) and Fischbacher and Gächter (2005) classify their subjects according to their contribution function. A subject is called a 'free rider' if and only if he or she contributes zero in all 21 cases. A subject is called a 'conditional cooperator' if the contribution schedule is a clearly positive function of the others' average contribution. A somewhat peculiar type is the 'triangle contributor' whose contribution is increasing in the others' contributions for low values and decreasing for high others' contributions. Figure 2 illustrates the (average) contribution function of the different types in the experiments by Fischbacher and Gächter (2005).

More than half of all subjects are 'conditional cooperators'. Twenty-three percent are 'free riders'. The rest are either 'triangle contributors', or non-classifiable 'others'. Fischbacher et al. (2001), and Gächter, Herrmann and Thöni (2005), who replicated this experiment in Russia, got very similar distribution of types and even of average contribution patterns.

Further studies that find evidence for heterogeneous cooperation preferences are, e.g., Ockenfels (1999), Kurzban and Houser (2005), and Burlando and Guala (2005).

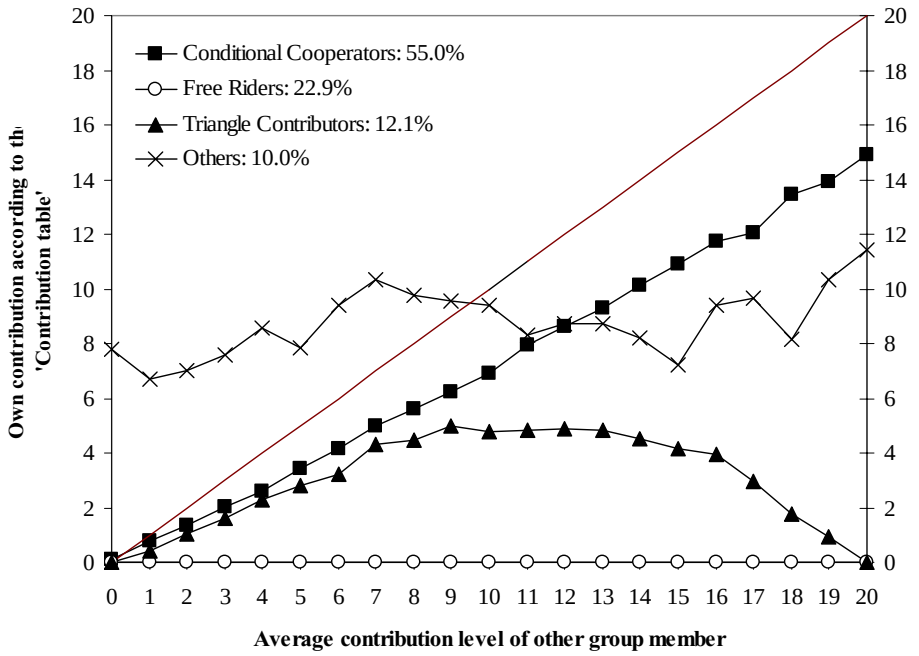

Figure 2: Average contribution function of types *Free Rider*, *Conditional Co-operator*, *Triangle Contributor*, and *Others*. Observations on the diagonal would correspond to the type of a perfect (i.e. one-to-one) *Conditional Co-operator*. Source: Fischbacher and Gächter (2005).

In summary, the evidence from the laboratory unambiguously shows that there is much more cooperation that is predicted by standard theory. Moreover, we find strong evidence that many people's attitude toward voluntary cooperation is conditional on other people's cooperation. This suggests that "warm glow" is not a dominant motivation. Moreover, the fact that many people contribute more the more others contribute also speaks against pure altruism explanations, because they predict that people reduce their own contributions when informed that others already contribute to the public good.

A second important finding is that people's contribution preferences are heterogeneous. While a large number of people seem to be conditional cooperators, a significant fraction of subjects is best characterized as free riders. Some others show more complicated patterns. In Section 3 I will discuss experiments that test directly for implications of preference heterogeneity. Yet, before I do so, I discuss evidence from the field.

2.2 *Evidence from field experiments*

Field experiments offer a great opportunity to test the behavioral relevance of laboratory findings in naturally occurring contexts (see also Harrison and List 2004). In this section I discuss a few field experiments that present results that are consistent with the lab evidence.

A first interesting study is by Frey and Meier (2004). Their subjects are students of the University of Zurich. Upon registering, each semester each student is asked whether, in addition to the tuition fee, he or should would like to make a contribution to two funds – one that helps needy students with cheap loans, and one that supports foreign students. A donation to the loans fund costs CHF 7 and CHF 5 to the support fund for foreign students. Students can either donate these fixed amounts or not. Intermediate donations are not possible. The data set comprises 37,624 students. For the field experiment, 2,500 non-freshmen students were randomly selected; 2000 of them received information about what others did. One thousand students received the information that a high fraction of others (64 percent) made a donation in the past; the remaining 1000 students got the information that a relatively low fraction (46 percent) made a donation in the past.[2] From 500 students expectations were elicited about the fraction of students who make a donation.

The results are consistent with theories of conditional cooperation. First, students who expect more others to donate are more likely to donate. The correlation between expressed expectations and actual donation is 0.34 ($p < 0.001$). Second, a logit analysis shows that those students who received the information that 64 percent of others has donated in the past, are more likely to donate than those who received the information that only 46 percent donated.

Croson (2005) conducts a field experiment on donations to a public radio station, which is a naturally occurring public good. The study is similar in spirit to Frey and Meier (2004). In a fund raising drive, people who call in to make a donation (for renewing their membership) are confronted with what others have donated in the past. Specifically, in the experimental condition (but not in the control condition) the experimenter read the following sentence: "We had another member, they contributed $75 [$180 or $300]", and right after that "How much would you like to pledge today?" Then they can make their pledge (any amount they wish). In total, 538 members called to make a donation. The benchmark for donation decision is the fund drive in the previous year, in which the average (median) amount donated was $135 ($75). The amounts used as the treatments correspond to the 50th percentile ($75), the 85th percentile ($180) and the 90th percentile ($300) in the previous fund drive. The results again support conditional cooperation. Callers who were confronted with a previous pledge of $300

[2] No deception was involved because real frequencies (resulting from different time periods) were used.

by another member donated significantly more than people in the control condition who were not confronted with that information; callers who received the $75 or $180 information, respectively, contributed more as well than the control group, but this effect is not significant.

In summary, the results from two field experiments support the importance of conditional cooperation in the field. In the next section I quickly discuss a study by Benz and Meier (2005) that tests to what extent the same person behaves conditionally cooperatively inside and outside of the lab. This is an important question, because lab experiments are sometimes criticized for their lack of "external validity".

2.3 *Connections between the lab and the field*

To gather information about the connection between lab and field behavior, the subjects in Benz and Meier (2005) take part in a lab experiment, where they make a donation decision. The same subjects are observed in a naturally occurring environment – the donation decisions to two student support funds described above and analyzed in Frey and Meier (2005). In one experiment (n=99), called 'social funds', the donation decision is to exactly the same funds as in the naturally occurring situation; in a second experiment (n=83), called 'charities', the donation decision is to another charity that is unrelated to university.

The results show that lab and naturally occurring behavior are correlated. In the 'social funds' experiment, the correlation between the average donation in the experiment and the average donation in the past four semesters is 0.28 ($p<0.01$). In the 'charities' experiment the correlation is very similar (0.27; $p<0.01$). A more refined statistical analysis that controls for socio-demographic variables in a multivariate regression analysis supports the main findings. Thus, although the lab is an artificial environment, one can observe behavior that is relevant in a naturally occurring environment.

In my view, this result strongly underscores the complementarity between the lab and the field. The lab is an "artificial" environment, whereas the naturally occurring situation is a context-rich environment. Depending on the research question, context-richness and artificiality are either a drawback or an advantage. The lab has the advantage that we can observe motivations and behavioral patterns in a degree of clarity that is most often not feasible outside the lab. The fact that we have observed conditional cooperation in tightly controlled lab experiments supports the interpretation of the field results as coming from conditional cooperation. The other side of the coin is that the field observation of conditional cooperation is important, because is tells us that it is not just a lab phenomenon.

In the following section I will use the power of the lab to rigorously test four implications of conditional cooperation and preference heterogeneity. I see these experiments as four behavioral models that might help us interpreting naturally occurring field situations in policy-relevant domains like tax morale and welfare state policies, but also in managerial domains like work-place behavior. The four models will also help me in guiding my discussion of consequences for public policy and management.

3. Four consequences of conditional cooperation and preference heterogeneity

In this section I will present four experiments that test four implications of conditional cooperation and preference heterogeneity in general. The testable consequences are that (i) in groups where group members are randomly selected *voluntary cooperation is bound to be fragile*; (ii) *there are group interaction effects,* i.e., people adapt their cooperation behavior to the relevant group they are a member of; (iii) *group composition matters*, i.e., in groups that are composed of 'like-minded types' (groups composed of either co-operators or free riders) we should see starkly different cooperation patterns; and (iv) *belief management matters*, i.e., factors that shift the belief about how much others contribute will influence contribution behavior. I discuss these four hypotheses and their experimental support in turn.

3.1 *Voluntary cooperation is fragile*

In this section I provide evidence that heterogeneous motivations in randomly composed groups will lead to fragile cooperation. The reason is that free riders presumably do not contribute to the public good while the conditional cooperators' contributions might be non-minimal, depending on their belief about other group members' contributions. During the repeated interaction the subjects learn the contributions of the other team members. The free riders have no reason to react to that information. The conditional cooperators on the other hand will update their beliefs. Given that the average conditional cooperator does not fully match the others' contribution the reaction will most likely be a decrease of contributions. There is no reason to expect that the remaining types (triangle contributors and 'others') will behave in a way that offsets the negative trend.

To rigorously test this argument, Fischbacher and Gächter (2005) combined the elicitation of contribution functions described above with a standard ten-period public goods game

played in the 'stranger' mode, i.e., in every period the groups of four are formed randomly out of all subjects in a session. As predicted, contributions actually fall over time (from initially 40 percent to ten percent by the last period).

Is this decline really due to the interaction of heterogeneously motivated types? Stringent support for the conjecture comes from using the elicited contribution functions for predicting contributions. Recall that the strategies asked subjects to indicate how much they are prepared to contribute to the public good for all feasible average contribution levels of the other group members. In the standard ten-period public goods game Fischbacher and Gächter (2005) also elicited in each period each subject's *belief* about the other group members' contributions. Therefore, we can – given a stated belief about other group members' average contribution – *predict* what a subject should contribute to the public good if he or she would be perfectly consistent with his or her elicited contribution function. Figure 3 depicts the actual average contributions in the ten rounds of the public goods game and the predicted contributions as a result of stated beliefs and contribution schedules.

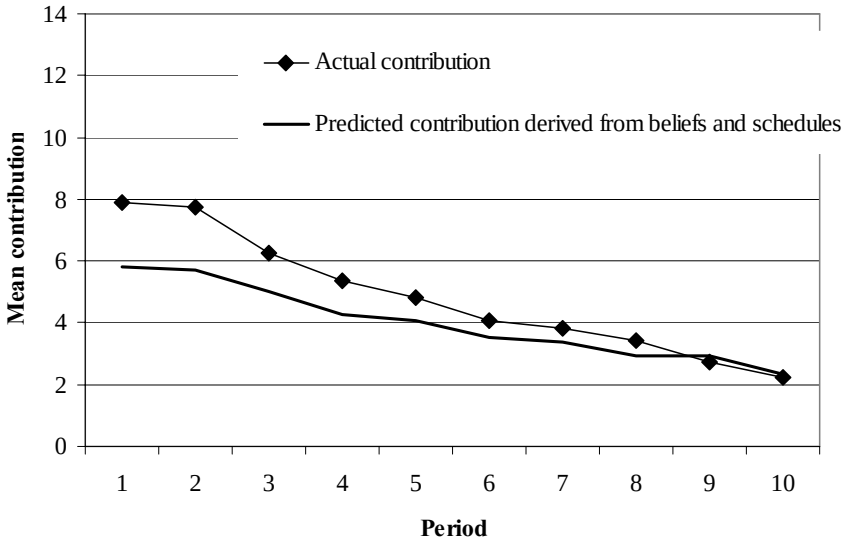

Figure 3: Average actual contributions and predicted contributions. Source: Fischbacher and Gächter (2005).

Although, compared with actual contributions, average predicted contributions are too low we find that predicted contributions, which are derived from the contribution functions and the elicited beliefs, actually decline and converge to the actual pattern. Therefore, this result supports the argument that preference heterogeneity leads to unstable cooperation.

3.2 *There are social interaction effects in cooperation*

If people are motivated by conditional cooperation, this may give rise to a "social interaction effect", which occurs if an individual changes his or her behavior as a *function of his or her respective group members' behavior*. Identifying social or group interaction effects (often also called 'neighborhood' or 'peer effects') is notoriously difficult (Manski 2000). The ideal data set would observe the same individual at the same time in different groups, which are identical – apart from different group members. Obviously, this is impossible in the field. By contrast, in the lab it is possible to come very close to this 'counterfactual state'. In an experiment, one is able to *observe decisions of the same subject at the same time* in two economically identical environments. The only reason to behave differently in these two environments is social interactions, i.e., the fact that a person is systematically affected by the behavior of his or her group members in the two environments. Falk, Fischbacher and Gächter (2005) test this idea in a design where every subject simultaneously is a member of two groups, group 1 and group 2, that provide two independent public goods. The two groups are identical except that the other group members in both groups are different people. Group composition stays constant for the twenty periods of the game. Falk et al. (2005) speak of a social interaction effect if the following holds: the larger the *difference* in contributions of group members in group 1 and group 2 in the previous period, the larger is the *difference* in current contributions of a group member to the two groups. Figure 4 provides the evidence.

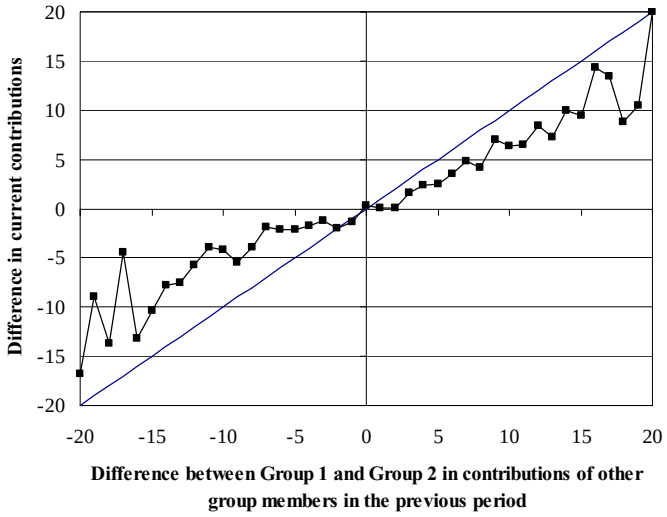

Figure 4: Social interaction effects: Difference in own contribution as a function of the group members' contributions in the two groups. Source: Falk, Fischbacher and Gächter (2005).

The results provide unambiguous support for the social interaction hypothesis. In a given period a majority of subjects contributes more to the group that has contributed more in the previous period. This result holds for all fourteen independent units of observations, which is a result that is very unlikely due to chance ($p<0.00007$).

3.3 *Group composition matters*

We have seen that a mixture of conditional cooperators and free riders is unfavorable for reaching cooperation in the public goods game. According to our third conjecture, conditional cooperators would presumably prefer to play the game with like-minded cooperators. Cooperation should be easy if the team players know that they are among like-minded group members. Similarly, if the "true game" subjects are playing is a game where cooperation is one of the equilibria (free riding being another one), then knowing that others are like-minded co-operators should make it easy for subjects to coordinate on cooperation and to prevent free riding. Likewise, if free rider types would know that they are among other free riders, free riding should be paramount.

Gächter and Thöni (2005) conducted an experiment where the subjects play in groups of 'like-minded' people. Thereby, like-mindedness refers to the type of a subject according to a classification whether one is a free rider or a conditional cooperator. The experiment starts with a one-shot public goods game. When all subjects have chosen their contribution the subjects are ranked according to their contribution. Then the subjects are reassigned to new groups. The reassignment works as follows. The three subjects with the highest contribution in the one-shot public goods game constitute a first group. The subjects with the fourth to sixth highest contribution are in the second group and so on. Finally, the three least cooperative subjects find themselves in the last group. The subjects are informed about the reassignment procedure only after they finished the first game. Then the subjects learn the contributions their new group members chose in the one-shot public goods game. In the new group subjects play a ten-period public goods game. It is also important to note that the subjects do not know the reassignment mechanism when choosing their contribution in the one-shot public goods game. Therefore, a high contribution in this game credibly reveals a cooperative attitude.

How do subjects play the public goods game when they know they are among like-minded people? Gächter and Thöni report the results from 18 groups of three subjects. The left panel of Figure 5 shows the results of the main treatment. In this game the maximal contribution is 20. For expositional ease the groups are divided into three classes (TOP, MIDDLE and

LOW) according to their average contribution in the one-shot public goods game. The three graphs show the average contribution during the ten periods separated by class. The unconnected dots in period zero show the average contribution in the one-shot public goods game that determines the group composition. The classes remain clearly separated over all periods. The groups in the TOP class consist to a large degree of subjects who contributed their entire endowment in the one-shot public goods game. These groups manage to maintain almost full cooperation until the penultimate period. The contributions of the MIDDLE class (consisting of subjects with intermediate contributions in the one-shot public goods game) show a similar pattern on a somewhat lower level. Surprisingly, also the subjects in the LOW class, who almost all chose a contribution of zero in the one-shot public goods game, manage to reach a certain level of cooperation in the repeated game.

The right panel of Figure 5 shows the results from a control experiment. In this experiment, groups are formed randomly as usual, i.e., there is no reassignment according to cooperativeness. In order to make the two treatments comparable the data is still separated into the three classes. However, the separation now merely reflects the fact that there is variance in the contributions.

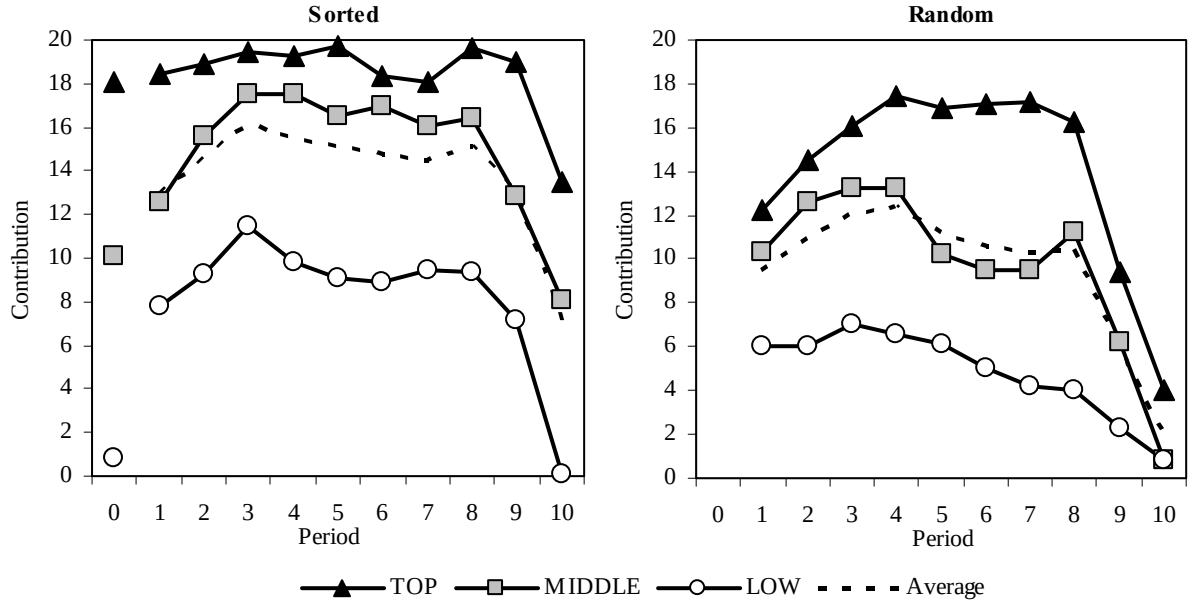

Figure 5, left panel: Average contributions over the ten periods for the TOP, MIDDLE and LOW class in the *Sorted* treatment. The unconnected dots in period zero are the average contributions in the *Ranking* treatment. Right panel: Average contribution of the most, intermediate and least cooperative groups over the ten periods. Source: Gächter and Thöni (2005).

What does the comparison between the left and the right panels of Figure 5 tell us about the effect of grouping like-minded subjects? First of all, cooperation in the TOP class of the sorted treatment is much higher than the average contribution in the random treatment (dotted

line in the right panel). However, the real value of the sorting mechanism becomes clear if we compare the TOP class with the most cooperative third of the groups in the random treatment. The average contribution of the TOP class of like-minded groups is significantly higher than the average contribution of the most cooperative third of the groups in the random treatment. The fact that even the groups in the LOW class contribute somewhat more if they know they are among like-minded people is surprising at first sight. However, if uncooperative subjects know that they are among themselves then it is clear that there are no cooperative subjects to free ride on. This presumably motivates even uncooperative subjects to contribute some of their endowment in order to encourage the other free riders to contribute as well. Yet, by the final period the contributions of these subjects drop to zero.

3.4 *Belief management matters*

Since the belief about others' contribution is important for conditional cooperators, our fourth conjecture says that any factor that moves these beliefs will influence cooperation. In the experiments of Fischbacher and Gächter (2005), for instance, beliefs evolved endogenously and mimicked the decline in cooperation. To test how beliefs can be influenced, Gächter and Renner (2005) developed a leader-follower design in a group of four players who stayed together for ten rounds, which was known to the subjects. Specifically, one group member was assigned 'leader'. All group members had the same payoff function (1). The sole difference between the leader and the followers was that the leader made the first contribution decision. This was observed by the followers who decided simultaneously about their contribution. Gächter and Renner (2005) also elicited the followers' beliefs about the contribution of the other followers. This allows them to determine how the leader's contribution influences the beliefs about other followers' contribution.

The line with the open squares in the left panel of Figure 6 shows that the leader's contribution in the first period positively influences the follower's beliefs about other follower's contributions. The first period is particularly interesting, because the followers have not yet made any observation about the other followers' actual contributions. The more the leader contributes in the first period, the higher is the followers' beliefs about what other follower will contribute. This is the main and most direct evidence that a leader "manages the followers' beliefs". In their actual contributions followers match their beliefs quite closely (see the line with the filled squares).

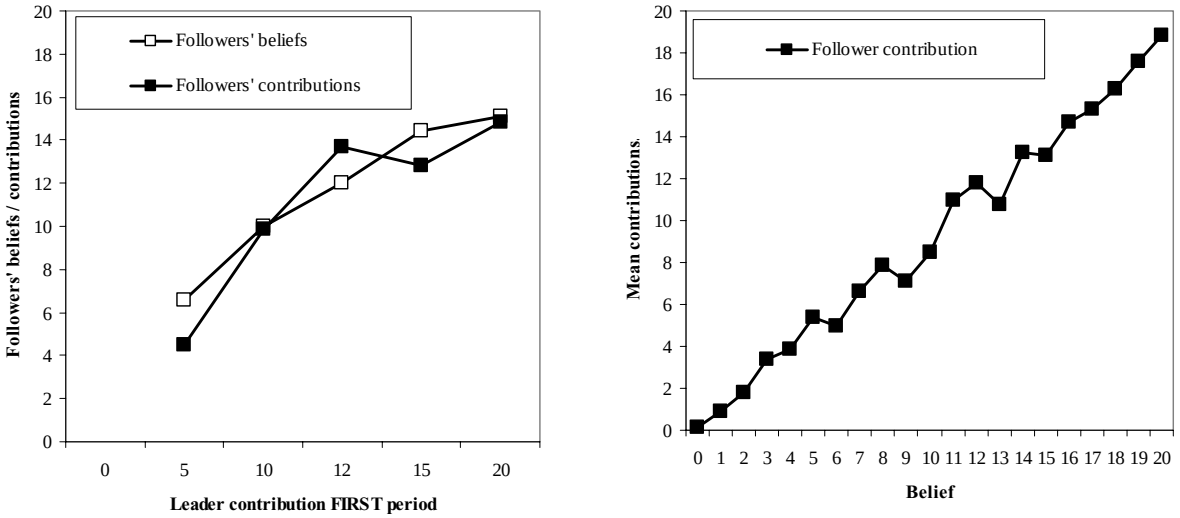

Figure 6: left panel: leader's contribution in the first period and followers beliefs and actual contributions in the first period; right panel: relationship between beliefs and followers' actual contributions over all rounds. Source: Gächter and Renner (2005).

Using the data from all periods, the right panel of Figure 6 shows that followers' beliefs and actual contributions are highly positively correlated. An econometric analysis reveals that these beliefs result from two sources: In a given period t>1, beliefs are highly significantly positively correlated with the leader's contribution in this period. Yet, beliefs are also highly significantly positively correlated with what the other followers contributed in the *previous* period (t–1). Moreover, quantitatively, the followers' contributions in t–1 are more important than the leader's contribution for the followers' beliefs about other followers' contributions in period t. Thus, there is an important path dependency in contributions. If the leader contributed little in the first period, followers are likely to contribute a small amount as well. This observation will – in addition to the leader's contribution – shape beliefs about other followers' contribution. In turn beliefs are – as the right panel of Figure 6 shows –positively correlated with actual contributions. In other words, a bad start will make it very hard for the leader to lead his group by good example to high contribution levels. By contrast, a bold leader who sets a good example right from the beginning will positively influence follower's beliefs and contributions.

In summary, in this section I have presented four experiments that have tested four implications of conditional cooperation and preference heterogeneity in general. As discussed above, I see these experiments as behavioral models that reveal something of the behavioral logic of conditional cooperation and preference heterogeneity. In the final two sections I will

therefore use these behavioral models to look at field phenomena[3] and to discuss implications for public policy and management.

4. Understanding field phenomena

4.1 *Charitable giving*

As a reaction to the humanitarian disaster in former Yugoslavia ten years ago three Austrian (charity) organizations set up the fund raising campaign "Nachbar in Not" to finance food, clothes and medical aid for the victims of the war. During three years people donated more than 950 millions Austrian Schillings (approx. €70 million) to "Nachbar in Not" alone - donations to other charity organizations not included. Another example of a very successful and very large charitable fund raising campaign that runs since many years around Christmas is "Licht ins Dunkel" by the Austrian broadcasting corporation (ORF).

Except for the huge donations these two examples have something else in common. In both campaigns it was practice to list the names, hometowns and donated amount of *all* donators either on television or in newspapers which supported the campaigns. Donations by well-known politicians and celebrities were particularly prominently featured. The results from the field experiments by Falk and Meier (2004) and Croson (2005) and the lab results on how leader contributions can shape followers' contributions suggests that fund raising organizers did not only rely on people's feelings of altruism, compassion and warm glow but also on conditional cooperation. Seed money effects are a related phenomenon that at least in part exploits the psychology of conditional cooperation (List and Lucking-Reiley 2002). Likewise, in fund drives, fundraisers, including charities, often make a symbolic gift to the donator. Reciprocity as a form of conditional cooperation predicts that nicer gifts will lead to higher donations. Falk (2004) tests this prediction in a field experiment and finds it unambiguously supported.

4.2 *Tax morale, benefit fraud, and corruption*

Norms of reciprocity and conditional cooperation might also influence tax morale. Controlling for detection probabilities conditional cooperators will be more likely to evade taxes or falsely claim welfare benefits if they have the impression that many others do the same. Too many cheaters can spoil tax morale. The evidence is consistent with this prediction. People are more likely not to cheat on their taxes if others honestly pay theirs (e.g., Andreoni et al. 1998; Rothstein 2000).

[3] See Falk (2003) and Fehr and Fischbacher (2002) for related discussions.

The prevalence of corruption also seems to be influenced by motivations similar to conditional cooperation (see Abbink et al. 2002 for an experiment and further references to the literature). There are also important social interaction effects in these phenomena (Bertrand et al. 2000; van der Klaauw and van Ours 2003), which is also predicted by conditional cooperation and our model of Section 3.2.

A particularly interesting observation is that the perception of the fairness of the tax system matters (Seidl and Traub 2001). Likewise, the treatment by authorities apparently is an important determinant for people's tax morale (Pommerehne and Weck-Hannemann 1996; Frey 1997; Goette and Kucher 1998; Scholz and Lubell 1998; Feld and Frey 2002). How can this be explained by our models? First, there may be a direct effect by the concerned individual who may reciprocate unfair treatment by authorities and/or the tax system by lower tax morale, simply because the taxpayer resents unfair treatment (Smith 1992). Second, there may be an indirect effect, via the beliefs about other tax payers' behavior. The reason is that if many people share similar feelings and experiences, then this will lower the belief that others have high tax morale, which further undermines tax morale. Similarly the government's trust in the honesty of its citizens, may lead to a direct effect of "trust breeds trust" (Feld and Frey 2002), presumably because people like to be considered trustworthy. Again, if such feelings are widespread, they may shape beliefs about other citizen's tax morale and hence reinforce the tax payer's tax morale.

A further interesting observation is that tax evasion at the Swiss cantonal level is higher in cantons where citizens have more direct democratic rights. According to our models, direct-democratic procedures may positively influence tax morale because they may affect the beliefs about others tax morale once a tax law is passed in a referendum. A referendum signals people's opinion about a topic, and the dissemination of opinions via the result of a referendum may shape people's beliefs about others' behavior. Tyran and Feld (2002) test this intuition in an experiment and find support for it.

4.3 *Solidarity and support for the welfare state*

Observers of welfare state policies (e.g., Wax 2000; Fong 2001; Fong et al. 2002; Lindbeck et al. 1999) point out that many people hold reciprocity norms that are akin to the conditional cooperation observed in our experiments. Fong et al. (2002) even argue that "people support the welfare state because it conforms to deeply held norms of reciprocity and conditional obligations to others". There is evidence that people resent certain welfare policies if they think that the recipient is a free rider who could earn his or her own living

(Wax 2000; Fong et al. 2002). Besley and Coate (1992, p. 175) in a paper on tax payer resentment (i.e., the resentment to finance welfare payments) quote a notable British columnist, Lynda-Lee Porter, which neatly expresses the psychology of such resentment: "Our bronzed, healthy, young hedonistic army of self-unemployed are holidaying by the sea at our expense this year and, yes I do resent it. I resent working to support the idle loafers who have a laugh at our expensively generous system which allows them to get away with legalised plunder."

4.4 *Work morale*

Our models predict that work morale is strongly shaped by the behavior of management and co-workers. First, there may be social interaction effects in that people adapt their work morale to those of their peers. Empirical evidence supports this prediction (Ichino and Maggi 2000; Falk and Ichino 2004).

Second, our leadership model, discussed in Section 3.4, suggests that managers may strongly influence morale. To our knowledge, there is no systematic evidence available, but some telling anecdotal evidence supports the point. For instance, in the wake of the Enron scandal, Lawrence Weinstein, the Head of Unisys, said about ethical standards in companies: "Once you as a CEO go over the line, then people think it's okay to go over the line themselves."[4] Thus, this quote clearly expresses the conviction that leading by example matters for the ethical behavior of employees. Moreover, our results from section 3.4 suggest that the CEO's behavior may have long-lasting consequences on company morale and culture because of path dependency effects.

Third, our result from section 3.3 that group composition matters may explain why companies sometimes fire workers, despite that firing looks like a policy of management by threats. Yet, Bewley (1999) notes that companies fire shirkers and incompetents to reestablish the work morale of the rest. This can be explained by our models. Recall that the experimental findings reported in section 3.1 and 3.3 suggest that in heterogeneous groups, contributions decline to low levels because the conditional cooperators stop cooperating once they experience free riding. If conditional cooperators know that they are among 'like-minded' cooperators, cooperation can be established at very high levels. In a company context, this may mean that work morale is undermined by even a few shirkers. Motivated workers may prefer that 'bad apples' are fired because they don't like being 'suckered' by their colleagues and because it reestablishes beliefs about others' team-spirit.

[4] Quoted after The Economist, July 27, 2002, p.58.

5. Consequences for public policy and management

In this section I briefly discuss policy implications that follow from the experimental findings and the four behavioral models that I have discussed above. I first look at implications for public policy (section 5.1) and then at consequences for management (section 5.2).

5.1 *Public policy*

Public policy is relevant mainly in the domains that I covered in the examples of sections 4.2 and 4.3. A first observation is that behavior by leaders – politician's and top officials' – may matter strongly for the morale of the citizens. Among other things, leaders are "belief managers". As the experiments of section 4.4 have shown, leading by example strongly shapes beliefs about what others are doing. Therefore, there is a 'multiplier effect', because a bad example (dishonesty in tax matters, corruption, and unethical behavior in other domains) may not only have direct effects on the concerned individual but also indirect belief effects about how others will react. Moreover, there may be strong path dependency effects, which may adversely affect morale in the long-run. Thus, leaders should be role models for which higher moral standards should hold than for normal citizens. In particular, leaders should be forced to resign quickly if there is confirmed evidence of dishonesty and inappropriate behavior.

Belief management happens not only through leaders, but also through effects like the perceived fairness of the tax system, fair treatment by authorities, and direct-democratic participation rights. The experimental results discussed above suggest that these factors are very important and should be strengthened. Tax reforms should improve the fairness of the tax system (based on careful evidence on how fairly the tax system is perceived) not only because fairness is desirable in its own right, but also because of its indirect effect on the beliefs about other citizen's tax morale. A similar conclusion holds for the reform of tax authorities. Presumably, the strongest effect on tax morale may come through direct-democratic participation rights (see, e.g., Feld and Frey 2002, and Tyran and Feld 2002), yet this is probably the most difficult to implement in representative democracies.

The experimental results from sections 3.1 and 3.3 suggest that free riders are the most important trigger for reduced cooperation. In the absence of punishment of free riders cooperation unravels, because the conditional cooperators reduce their cooperation as well. Experiments have shown that this result can be overturned if punishment of free riders is

possible (e.g., Fehr and Gächter 2000) or if the free riders are excluded from the group (Gächter and Thöni 2005 – see section 3.3; Cinyabuguma et al. 2005). Therefore, policy should aim to punish free riding (i.e., tax evasion, benefit fraud, and corruption). The experiments described above suggest that the goal should be to punish the free riders and at the same time to maintain the optimistic beliefs of the cooperators, by reassuring them that they will not be 'suckered' by the free riders, so that they continue to uphold their morale together with other 'like-minded cooperators'.

Yet, apart from the legal implementation (which might be relatively simple) this is no easy task at all, given the behavioral regularities discussed above. The reason is that punishment may entail monitoring and a general distrust of the citizens. This is problematic for two reasons. First, there is evidence that monitoring may crowd out intrinsic motivation and reciprocal behavior (Frey 1993; Bohnet, Frey and Huck 2001; Fehr and Gächter 2002). Second, monitoring may express distrust, which, in addition to the crowding out effect, may have detrimental effects on the beliefs about the tax morale of other tax payers. Thus, in order to avoid the negative side effects of distrusting most citizens, policies should aim at punishing the big offenders severely and treat the mild offenders (provided they are no 'serial offenders') mildly (by not using the full force of penal law, for instance). This has two advantages. First, strong sanctions have a deterrence effect, and they also reassure the honest citizens that large-scale anti-social behavior will be punished, which reduces the 'sucker effect'. Second, by trusting citizens and by fostering the fairness of the tax system and the tax authorities, crowding out effects of intrinsic motivation and voluntary cooperation may be avoided.

5.2 *Management*

The conclusions for management are very similar. First, leaders should be aware that they are role models who set an example and may strongly shape corporate cultures for reasons of path dependency in behaviors. Therefore, like politicians, they should be held to high ethical standards.

Second, the problem of punishment of shirkers in an organization is also similar to the problems of how to treat anti-social behavior in the public policy domain. Management by threats will not create loyalty and may undermine intrinsic motivation and voluntary cooperation. Therefore, firing shirkers according to procedurally fair standards (see Benz 2005) may help maintaining high work morale among the team-spirited workforce.

Third, since group composition effects matter strongly for cooperative behavior hiring of team-spirited people is crucial if teamwork is important on the job. Composing teams of like-minded team players can help maintaining high cooperation levels without any threat and negative side effects of monitoring and distrust.

6. Summary

In this paper I have discussed experimental evidence from the lab and the field that many people's attitude to voluntary cooperation is characterized by conditional cooperation. I believe that conditional cooperation helps us better understanding important phenomena in the field, like tax morale and attitudes toward the welfare state. Since beliefs about others' behavior are highly relevant for voluntary cooperation if most people are conditional cooperators, policy should not only take into account the incentive effects on the behavior of an individual, but also how policy affects the beliefs and behavior of the presumed majority of citizens who are conditional cooperators.

References

Abbink, K., Irlenbusch, B., Renner, E., 2002. An experimental bribery game. Journal of Law, Economics and Organization 18(2), 428-454.

Anderson, S.P., Goeree, J.K., Holt, C., 1998. A theoretical analysis of altruism and decision error in public goods games. Journal of Public Economics 70, 297–323.

Andreoni, J., 1988. Why Free Ride? Strategies and Learning in Public Goods Experiments. Journal of Public Economics 37, 291-304.

Andreoni, J., 1990. Impure altruism and donations to public goods: A theory of warm-glow giving. Economic Journal 100, 464–477.

Andreoni, J., and Croson, R., 1998. Partners versus Strangers: Random Rematching in Public Goods Experiments, forthcoming in Handbook of Experimental Economic Results, ed. by C. Plott and V. Smith.

Andreoni, J., Erard, B., Feinstein, J., 1998. Tax compliance. Journal of Economic Literature XXXVI, 818-860.

Benz, M. 2005. The relevance of procedural utility for economics. Mimeo, University of Zurich.

Benz, M., Meier, S., 2005. Do people behave in experiments as in real life? Evidence from donations. Mimeo, University of Zurich.

Bertrand, M., Luttmer, E., Mullainathan, S., 2000. Network Effects and Welfare Cultures. Quarterly Journal of Economics 115(3), 1019-1055.

Besley, T., Coate, S. 1992. Understanding taxpayer resentment. Journal of Public Economics 48, 165-183.

Bewley, T. 1999. Why don't wages fall in a recession? Harvard University Press.

Bohnet, I., Frey, B.S., Huck, S., 2001. More Order with Less Law: On Contract Enforcement, Trust, and Crowding. American Political Science Review 95, 131-144.

Burlando, Roberto and Francesco Guala, 2005. Heterogeneous Agents in Public Goods Experiments. Experimental Economics 8(1), 35-54.

Camerer, C., 2003. Behavioral game theory. Princeton: Princeton University Press.

Cinyabuguma, M., Page, T., Putterman, L., 2005. Cooperation under the threat of expulsion in a public goods experiment. Journal of Public Economics 89, 1421-1435.

Croson, R., 2002. Theories of altruism and reciprocity: evidence from linear public goods games. University of Pennsylvania, Mimeo.

Croson, R. 2005. Field experiments in charitable contribution: The impact of social influence on the voluntary provision of public goods. Mimeo, Wharton University.

Dawes, R. 1980. Social dilemmas. Annual Review of Psychology 31, 169-193.

Falk, A. 2003. Homo Oeconomicus versus Homo Reciprocans: Ansätze für ein neues Wirtschaftspolitisches Leitbild? Perspektiven der Wirtschaftspolitik 4(1), 141-172.

Falk, A. 2004. Charitable giving as a gift exchange. Evidence from a field experiment. Discussion Paper IZA-DP 1148, Bonn.

Falk, A., Fischbacher, U., Gächter, S. 2005. 'Living in two neighborhoods' – Social interaction in the lab. Working paper No. 150, Institute for Empirical Research in Economics, University of Zurich.

Falk, A., Ichino, A. 2004. Clean evidence on peer pressure. Discussion Paper IZA-DP 732, Bonn.

Fehr, E., Fischbacher, U. 2002. Why social preferences matter- the impact of non-selfish motives on competition, cooperation and incentives. Economic Journal 112, C1-C33.

Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. American Economic Review 90(4), 980-994.

Fehr, E., Gächter, S., 2002. Do incentive contracts undermine voluntary cooperation? Working paper No. 34, Institute for Empirical Research in Economics, University of Zurich.

Fehr, E., Schmidt, K. 2003. Theories of fairness and reciprocity: Evidence and economic applications. In: Matthias Dewatripont, Lars Hansen & Stephen Turnovsky (Eds.). *Advances in Economics and Econometrics. Theory and Applications*. Vol. I. (pp. 208-257). Cambridge: Cambridge University Press.

Feld, L., Frey, B.S. 2002. Trust breeds trust: How taxpayers are treated. Economics of Governance 3, 87-99.

Feld, L., Tyran, J.-R. 2002. Tax evasion and voting: An experimental analysis. Kyklos 55(2), 197-222.

Fischbacher, U., 1999. z-Tree, Toolbox for readymade economic experiments: Experimenter's manual. Working paper No. 21, Institute for Empirical Research in Economics, University of Zürich.

Fischbacher, U., Gächter, S. 2005. Heterogeneous social preferences and the dynamics of free riding. Mimeo, University of Nottingham.

Fischbacher, U., Gächter S., Fehr E., 2001. Are people conditionally cooperative? Evidence from a Public Goods Experiment. Economics Letters 71, 397-404.

Fong, C., 2001. Social preferences, self-interest, and the demand for redistribution. Journal of Public Economics 82, 225-246.

Fong, C., Bowles, S. Gintis, H. 2004. Reciprocity, self-interest, and the welfare state. Forthcoming in: Ythier, J.M., Kolm, S. and Gerard-Varet L.-A., Handbook on the Economics of Giving, Reciprocity and Altruism. Amsterdam: Elsevier.

Frey, B.S., 1993. Shirking or work morale? The impact of regulating. European Economic Review 37, 1523-1532.

Frey, B.S., 1997. A constitution for knaves crowds out civic virtues. Economic Journal 107, 1043-1053.

Frey, B.S., Meier, S., 2004. Social comparisons and pro-social behavior. Testing 'conditional cooperation' in a field experiment. American Economic Review 94(5), 1717-1722.

Gächter, S., Herrmann, B., Thöni, C., 2003. Understanding Determinants of Social Capital. Cooperation and Informal Sanctions in a Cross-Societal Perspective. Mimeo, University of St. Gallen.

Gächter, S., Renner, E., 2005. Leading by example in the presence of free rider incentives. Mimeo, University of Nottingham.

Gächter, S., Thöni, C., 2005. Social learning and voluntary cooperation among like-minded people. Journal of the European Economic Association 3, 303-314.

Goette, L., Kucher, M., 1998. Trust Me. An Empirical Analysis of Taxpayer Honesty. Finanzarchiv 55, 429-444.

Harrison, G., List, J. 2004. Field experiments. Journal of Economic Literature XLII, 1009-1055.

Ichino, A., Maggi, G. 2000. Work environment and individual background: Explaining regional shirking differential in a large Italian firm. Quarterly Journal of Economics 115(3) 1057-1090.

Kagel, J., Roth, A.E. (Eds.), 1995, Handbook of Experimental Economics. Princeton: Princeton University Press.

Kahan, D. 1997. Social influence, social meaning, and deterrence. Virginia Law Review 83(2), 349-395.

Kelley, H. and Stahelski A., 1970. Social interaction basis of cooperators' and competitors' beliefs about others. Journal of Personality and Social Psychology 16, 190-197.

Keser, C., van Winden, F., 2000. Conditional cooperation and voluntary contributions to public goods. Scandinavian Journal of Economics 102(1), 23-39.

Kurzban, R., and Houser, D., 2005. An Experimental Investigation of Cooperative Types in Human Groups: A Complement to Evolutionary Theory and Simulations. *Proceedings of the National Academy of Sciences*, 102(5), pp. 1803-1807.

Ledyard, J., 1995. Public Goods: A survey of experimental research. In: Kagel, J. & Roth, A.E. (Eds.), Handbook of Experimental Economics. Princeton University Press, Princeton.

Lindbeck, A., Nyberg, S., Weibull, J. 1999. Social norms and economic incentives in the welfare state. Quarterly Journal of Economics CXIV(1), 1-35.

List, J., Lucking-Reiley, D. 2002. The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign. Journal of Political Economy 110(1), 215-233.

Manski, C., 2000. Economic Analysis of Social Interactions. Journal of Economic Perspectives, 14(3), 115-136.

Ockenfels, A., 1999. Fairness, Reziprozität und Eigennutz – ökonomische Theorie und experimentelle Evidenz. Die Einheit der Gesellschaftswissenschaften, Bd. 108, Mohr Siebeck, Tübingen.

Palfrey, T.R., Prisbrey, J.E., 1997. Anomalous behavior in public goods experiments: How much and why? American Economic Review 87, 829–846.

Pommerehne, W., Weck-Hannemann, H., 1996. Tax rates, tax administration and income tax evasion in Switzerland. Public Choice 88, 161-170.

Rothstein, B., 2000. Trust, social dilemmas and collective memories. Journal of Theoretical Politics 12(4), 447-501.

Samuelson, P., 1954. The pure theory of public expenditures. Review of Economics and Statistics 36(4), 387-389.

Scholz, J.T., Lubell , M., 1998. Trust and Taxpayers: Testing the Heuristic Approach to Collective Action. American Journal of Political Science 45, 160-178.

Seidl, C., Traub, S., 2001. Taxpayers' attitudes, behaviour, and perception of fairness. Pacific Economic Review 6(2), 255-267.

Smith, K.W., 1992. Reciprocity and fairness: Positive incentives for tax compliance. In: Slemrod, J. (ed.) Why People Pay Taxes: Tax Compliance and Enforcement. Ann Arbor, MI: University of Michigan Press.

Sugden, R., 1984. Reciprocity: the supply of public goods through voluntary contributions. Economic Journal 94, 772-787.

Tyran, J.-R., Sausgruber, R., forthcoming. A little fairness may induce a lot of redistribution in democracy. European Economic Review.

van der Klaauw, B., van Ours, J., 2003. From welfare to work: Does the neighborhood matter? Journal of Public Economics 87, 957-985.

Wax, A. 2000. Rethinking welfare rights: Reciprocity norms, reactive attitudes, and the political economy of welfare reforms. Law and Contemporary Problems 63, 257-298.