**Preprints of the**
**Max Planck Institute for**
**Research on Collective Goods**
**Bonn 2009/29**

# We Are Not Alone: The Impact of Externalities on Public Good Provision

Christoph Engel
Bettina Rockenbach

MAX PLANCK SOCIETY

Preprints of the
Max Planck Institute
for Research on Collective Goods Bonn 2009/29

# We Are Not Alone:
# The Impact of Externalities on Public Good Provision

Christoph Engel / Bettina Rockenbach

August 2009, revised December 2011

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
http://www.coll.mpg.de

# We Are Not Alone:

# The Impact of Externalities on Public Good Provision

**Christoph Engel and Bettina Rockenbach***

**December 1, 2011**

Abstract:

Public good provision is often local and also affects bystanders. Is provision harder if contributions harm bystanders, and is provision easier if outsiders gain a windfall profit? In an experiment we observe that both positive and negative externalities reduce provision levels whenever actors risk falling back behind bystanders. The mere presence of unaffected bystanders already dampens contributions. This behavior seems to result from the interplay of two motives: the desire to realize opportunities for joint gains, and concerns for comparative performance. Individual payoff comparisons to the other actors as well as to individual bystanders drive contributions down.

Keywords: Public Good, Externality, Conditional Cooperation, Inequity Aversion

JEL Classification: C91, C92, D03, D43, D62, H23, H41, L13

---

* We thank Armin Falk, Özgür Gürerk, Karim Sadrieh and seminar audiences at the University of Tilburg, the Gruter Institute, the University of Cologne, the University of Bonn, the University of Erfurt for valuable comments. Financial support by the Deutsche Forschungsgemeinschaft through RO3071 is gratefully acknowledged.

Authors' addresses:
Christoph Engel, Max Planck Institute for Collective Goods, Bonn and University of Bonn, Kurt-Schumacher-Str. 10, 53113 Bonn, Germany, e-mail engel@coll.mpg.de
Bettina Rockenbach, University of Cologne, Albertus Magnus Platz, 50923 Cologne, Germany, email: bettina.rockenbach@uni-koeln.de

# 1. Introduction

The essence of many social problems is the temptation to free ride on others' contributions to the provision of a public good. This essence has been backed by a rich experimental literature (Ledyard 1995; Fehr and Gächter 2000; van Dijk, Sonnemans et al. 2002; Masclet, Noussair et al. 2003; Page, Putterman et al. 2005; Potters, Sefton et al. 2005) and is corroborated in the field (Ostrom, Dietz et al. 2002; Anderson, Mellor et al. 2004; Andersen, Bulte et al. 2008). Behavioral research has shown that the core of the problem is not naked greed, but a hurt sense of fairness. In experimental populations many participants "conditionally cooperate". They are happy to make substantial contributions to a joint project as long as they believe a sufficient portion of others will do so as well (Keser and van Winden 2000; Brandts and Schram 2001; Fischbacher, Gächter et al. 2001; Frey and Meier 2004; Croson, Fatas et al. 2005; Fischbacher and Gächter 2010). The important news for policy makers is that it need not be necessary to force everyone to contribute. It may be enough to make sure that the risk of being the sucker is not too strong, or too salient.

Yet unfortunately political reality is often more complex. Public good provision is often local and also affects people outside the borders of the community. Equatorial countries preserving the rain forest do not only save their national ecosystems, but the world's climate and biodiversity along with it. If a metropolitan area subsidizes the opera house, it attracts visitors from further away who do not pay local taxes. In these examples public goods provision is not only domestically valuable but additionally creates a positive externality for outsiders. On the other hand, the successful provision of a (local) public good may create negative external effects. Take a country close to the source of an international river, building a dam to secure irrigation water and energy for its population. This deprives countries closer to the estuary of the river's benefits. Or, think of a municipality constructing a landfill close to its borders to keep garbage off its streets. This puts the groundwater in the neighboring community at risk. Of course, one of the economically most prominent examples for negative externalities of cooperation is the formation of a cartel. Successful cooperation among suppliers imposes damage on customers.

How does the existence of external effects affect public goods provision? Do positive external effects make provision "easier" while negative external effects decelerate the provision process? In this paper we tackle this question experimentally as well as theoretically. We model a linear public goods game with externalities on bystanders. In the *positive externality treatments* bystanders profit and in the *negative externality treatments* they suffer from the actors' contributions to the public good. In the *no externality treatment* bystanders are present, but their payoff is unaffected by the actors' provisions. The situation is asymmetric as bystanders have no direct means to influence the actors' payoffs. We not only vary the direction of the externality, but also vary the initial endowment of bystanders. That way, we are able to disentangle the effect of the direction of the externality from the effect of payoff differences between actors and bystanders.

Our experiments provide us with remarkable results. Already the mere presence of unaffected bystanders with the same endowment as actors substantially reduces contributions to the public good. Indeed, contributions are also reduced if bystanders have a higher endowment than actors, be that in a positive or in a negative externality case. If, however, in the positive externality case, bystanders are poorer than actors or, in the negative externality case, bystanders and actors have the same endowments, actors' contributions are not significantly different from the case of no bystanders.

We explain these observations with a combination of two motivating forces, neither of which would be sufficient in isolation to explain our data. Narrowly self-interested participants expecting a sufficient fraction of other active players to be cooperative may contribute in a finitely repeated game, but this cannot explain the observed treatment differences. Social preferences in the form of inequity aversion alone would only provide an explanation of the observed treatment differences if all actors shared an implausibly high aversion to outperforming others. We show that the interaction of repeated game effects and inequity aversion explains our observations. Actors cooperate the more, the more others cooperate, but additionally condition their contributions on the past difference between their own payoff and the payoff of the other actors as well as the payoff difference to the passive bystanders. If unaffected bystanders have the same endowment as actors, contributing actors risk receiving a lower payoff than bystanders. This risk is even more pronounced if bystanders have a higher endowment than actors, be that in a positive or in a negative externality case. This risk of falling behind bystanders significantly reduces actors' contributions in these cases and leads to significantly lower contributions than absent any bystanders. In comparative terms, however, poor bystanders in the positive externality case and equally endowed bystanders in the negative externality case are still worse off than actors making positive contributions. In these treatments, since negative payoff comparisons with respect to bystanders are not an issue, actors contribute as long as the others do so as well, and contributions are not different from the baseline of no bystanders.

Our results point to a limitation of self-governance. Conditional cooperators need institutions to protect them against the risk of being the sucker, especially with respect to outsiders gaining a windfall profit. By design, the institutional environment of our experiment did not provide this protection. In this light, going back to our examples, it becomes understandable why equatorial countries are compensated for preserving the rain forest by being exempted from the obligation to reduce $CO_2$ emissions; or why municipalities tax secondary residences, using the second home as a proxy for the benefit from local public goods.

One could think of even more general policy implications of these findings. As long as states would strictly maximize the aggregate utility of their citizens, many transnational public goods would be provided. Even if other states receive a windfall profit, the benefit for the nationals of the providing state would often still be large enough. Yet government has to defend higher taxes and onerous regulation vis-à-vis the citizenry. Not so rarely, political support for an otherwise sensible intervention falters if this gives outsiders a free lunch. A striking illustration is defense. Often, if one country disciplines a rogue state, many other countries benefit as well, yet save their soldiers' lives. And one sees why federations like the United States of America and confederations like the European Union have grown so large: under the federal umbrella, beneficiaries cannot so easily escape contributing their fair share.

From a policy perspective, our finding on the mere presence of bystanders is no less troublesome. If those who are asked to contribute run the risk of falling behind members of an external benchmark group, this aggravates the social dilemma. In political reality, in the short run benchmarks escape the control of policy makers. The media can always draw unfavorable comparisons. Yet in the long run, political action can change with whom citizens compare themselves.

Of course, all our examples are embedded in a much richer environment than the one we modeled in our experimental game. Yet in all examples, the underlying conflict has the structure of a public goods dilemma for the internals, and it invites potentially unfavorable comparisons with outsiders. If internal cooperation engenders a positive or negative externality, for actors their comparative position is not a given, but open to their action. Our results suggest that policy makers should be concerned that conditional cooperation is hampered when internal cooperation worsens the competitive position of actors, compared with the outsiders.

In section 2 we discuss the related experimental literature. In section 3 we introduce the game and in section 4 we embed our research question into the theoretical literature and derive hypotheses to be tested with our data in section 5. Section 6 concludes. The appendices provide supplementary material, including the instructions.

## 2. Related Literature

To the best of our knowledge no experimental study on public goods provision with externalities on inactive others has been conducted so far. Surprisingly, even in other contexts there are only a few studies which have aspects of externalities. Güth and van Damme (1998) present an ultimatum game with an externality on an inactive third player who has no say. The proposer offers how to divide the pie between three players. The division is executed if and only if the responder accepts. Otherwise, all three players receive nothing. The externality is the same in all treatments. If the responder only learns how much the proposer wants to allocate to the outsider, proposals are lowest. They are intermediate if the proposer only learns how much she gets in case she accepts. They are highest if the responder fully knows how the proposer wants to distribute the pie. Bolton and Ockenfels (2010) have an active player choose between a safe option and a lottery. Both affect an inactive outsider. If the safe option gives the active player a lower payoff than the inactive one, players choose the lottery more often. Abbink (2005) plays a two-person bribery game in which corruption negatively affects passive workers. He concludes that reciprocity between briber and official overrules concerns about distributive fairness towards other members of the society. Ellman and Pezanis-Christou (2010) study how a firm’s organizational structure influences ethical behavior towards passive outsiders. A firm of two players decides on its production strategy, which influences a passive third player. They find that horizontally organized firms in which the firm’s decision corresponds to the average of both individual decisions are less likely to harm the outsider than consensus-based firms or firms in which one of both members is the boss.

Studies with effects on active others are more common. Bornstein and colleagues extensively study team competition in various contexts (for an overview see Bornstein 2003) and find that in social dilemmas the competition with another group increases in-group cooperation. Abbink, Brandts et al. (2010) find that group members punish each other more severely if the group is in conflict with another group. The group position may be interpreted as a joint project of the group. Okada and Riedl (2005)**;** Kosfeld, Okada et al. (2009) study endogenous group formation in a public-goods setting. Players may declare their interest in an organization that demands full contributions of its members. Implementation of the organization is costly and requires the anonymous approval of all those interested. Non-members of the organization freely determine the size of their contribution to the public good. Both groups (members and non-members) contribute to a global public good. Thus, in contrast to our study, contributions of both groups have mutual positive effects and group association is endogenous, while we impose it. Kosfeld et al. find that in 70 to 100 percent of cases an organization was implemented in the final (of 20) rounds. Remarkably, around 75 percent of these organizations included *all* participants. If public goods are nested, such that simultaneously an inner and an outer group are affected, contribution patterns are sensitive to which dimension of the externality is made salient (Wit and Kerr 2002). If the marginal per capita rate for the global public good exceeds the marginal per capita rate for the local public good, participants contribute more to the global good, without reducing contributions to the local good (Blackwell and McKee 2003). Participants give more to the global good, the more a country is exposed to globalization (Buchan, Grimalda et al. 2009).

In an indirect way, the experimental literature on oligopoly also provides evidence. Collusion is significantly lower if the opposite market side is represented by real subjects (collusion rates of about 7%), rather than a computer bidding a predetermined demand function (collusion rates of about 43%) (cf. the meta-study by Engel 2007). This might indicate that participants shy away from imposing harm on other participants, which would imply that cooperation is lower if it entails a negative externality.

## 3. A Public Goods Game with Externalities

We introduce a linear public goods game in which public goods provision may cause externalities to non-actors. The game consists of $n_A > 0$ active players, the *actors*, and $n_B \geq 0$ passive players, the *bystanders*. Actors are endowed with $e_A$ and may contribute any amount $0 \leq g_i \leq e_A$ to a public good, which benefits all actors. As in a standard public goods game, the sum of all actors' contributions $G = \sum_{k=1}^{n_A} g_k$ is augmented by $a \cdot n_A$ and then equally distributed among the actors. The parameter $\frac{1}{n_A} < a < 1$ is the marginal per capita rate (MPCR) that specifies the marginal individual return each actor receives from her own contribution to the public good. The actors' payoff is given in equation (1):

$$(1) \qquad \pi_i^A = e_A - g_i + aG, \qquad i = 1,\ldots,n_A$$

Bystanders receive an endowment $e_B$ and cannot contribute to the public good. But – dependent on the parameter $b$ – they either benefit from ($b > 0$), suffer from ($b < 0$), or are unaffected by ($b = 0$) the contributions of the actors. Accordingly, for a given $b$ all bystanders earn an identical payoff which is solely determined by the actors' actions and is out of the bystanders' control. The profit function of bystanders is given by equation (2).

$$(2) \qquad \pi^B = e_B + bG$$

*Experimental Implementation*

In our experimental implementation we set the endowment of actors to be $e_A = 20$, and keep $a = 0.4$ fixed. For the sake of comparability we choose parameters that are standard in public goods experiments without externalities[1]. Our treatment subject groups consist of 7 subjects, $n_A = 4$ actors and $n_B = 3$ bystanders. We have chosen 3 bystanders so that the number of insiders remains larger (and therefore arguably more important), but that the number of bystanders is substantial. Our control subject group consists of $n_A = 4$ actors and no bystanders ($n_B = 0$). In all treatments, subjects play the above game repeatedly over 10 announced rounds in fixed groups. We restrict the contribution rates to be either $g = 0, g = 10,$ or $g = 20$. This limitation of the action space makes sure that upward or downward shifts from period to period are large. We can therefore be sure that changes in contributions are meaningful. We use changes in contributions to test our theoretical expectations about driving forces for treatment differences. Both active players and bystanders receive feedback about the sum of the contributions of all active players, their own period payoff and the period payoff of each bystander. From the instructions all participants know the actors' payoff function, and are thereby indirectly informed about the average performance of the (remaining) actors.

[1] In the meta-study by Zelmer (2003) on 27 studies with 711 distinct groups, the mean MCPR was 0.404 and the average number of periods was 10.4.

In our six treatments, we vary parameters in two dimensions. We first vary the way in which bystanders are influenced by the contributions of actors, i.e. we vary $b$. In the *positive externality treatments* $PE$, we set $b = 0.2$. In the *negative externality treatments* $NE$, we set $b = -0.2$. The choice of $b$ follows the same logic as the composition of the group. We want to study a case where the externality matters, but less so than the effect of contributions on insiders. In the *no externality treatment Nox7,* we set $b = 0$. Our second source of variation is the endowment bystanders receive upfront in every round. We implement symmetric and asymmetric endowments. In the symmetric treatments *Nox7, PE20, NE20*, bystanders have the same endowment of 20 tokens as have actors. In the interest of disentangling the effects caused by the direction of the externality and effects caused by payoff differences, we also vary bystander endowment. Treatment *PE0* gives bystanders no upfront endowment. In this treatment, active players can never fall behind bystanders. In the same spirit, treatment *NE60* makes bystanders so rich (with an endowment of 60) that active players have a lower payoff, however they perform. Whenever actors make positive contributions to the public good, bystanders are worse off in *NE20*. It therefore did not seem necessary to test a situation where bystanders have an endowment of zero and have to pay the experimenter in case of any cooperation. However, the opposite case is of interest. Is the willingness to contribute influenced if bystanders who were affluent in the first place get a windfall profit? This we test in treatment *PE40*. Finally, to have a proper benchmark, we compare all treatments to a standard voluntary contribution mechanism in a group of four, our control treatment *Nox4*. We thus compare all treatments with a baseline where bystanders are neither affected by internal cooperation, nor even present in the lab. Table 1 summarizes experimental parameters.

| treatment | number of actors $n_A$ | number of bystanders $n_B$ | Actor endowment $e_A$ | bystander endowment $e_B$ | actor mpcr $a$ | marginal effect on bystanders $b$ |
|---|---|---|---|---|---|---|
| *Nox4* | 4 | 0 | 20 | - | 0.4 | - |
| *Nox7* | 4 | 3 | 20 | 20 | 0.4 | 0 |
| *PE0* | 4 | 3 | 20 | 0 | 0.4 | 0.2 |
| *PE20* | 4 | 3 | 20 | 20 | 0.4 | 0.2 |
| *PE40* | 4 | 3 | 20 | 40 | 0.4 | 0.2 |
| *NE20* | 4 | 3 | 20 | 20 | 0.4 | -0.2 |
| *NE60* | 4 | 3 | 20 | 60 | 0.4 | -0.2 |

**Table 1**
**Treatments**

In treatment *Nox7* and in all positive externality treatments, contributions both serve the augmentation of the actors' profit and the joint payoff of actors and bystanders. Thus, joint payoff maximization demands full contributions of all actors, independent of whether the actors strive for maximizing the joint profit of actors alone or the joint profit of actors and bystanders. With negative externalities, in principle there is potential for a motivational conflict between augmenting the joint payoff of actors and reducing the joint payoff of all 7 participants. To rule out that conflict, and to have more scope for disentangling motives, we chose parameters that yield constant joint payoffs. In our negative externality settings actors' joint profit is $\Pi^A = 4 \cdot e_A + 0.6 \cdot G$ and bystanders' joint profit is $\Pi^B = 3 \cdot e_B - 0.6 \cdot G$. Thus, the net gain of actors is exactly identical to the net loss of bystanders. This choice of parameters makes sure that contributing actors cannot be motivated by efficiency. Whatever actors do for themselves is to the detriment of bystanders and neutral to the entire "society". This way, they cannot assuage bad feelings by an efficiency excuse.

*Conduct of the Experiment*

The experiment was run at the University of Erfurt (*elab*) with a computerized interaction using z-Tree (Fischbacher 2007). Subjects that never played a public goods experiment were invited using ORSEE (Greiner 2004). Each subject played in one of the seven parameter constellations (six treatments and control) and no subject played in more than one. We collected nine independent observations in each parameter constellation, adding up to 63 independent observations with a total of 414 subjects of various majors.[2]

## 4. Theoretical considerations and hypotheses

Narrowly self-interested actors, i.e. actors solely motivated by the maximization of their own monetary gains, are completely unaffected by the presence of bystanders and follow their dominant strategy of free-riding on the public good provision. Consequently, contributions of zero prescribe the unique Nash equilibrium of the stage game and, under common knowledge of rationality, the unique subgame perfect equilibrium of the finitely repeated game. There are

---

[2] Each session lasted about one hour and subjects earned on average 15.18 € in the control *Nox4*, 14.19 € in treatment *Nox7* (14.29 € for actors, and 14.06 € for bystanders), 11.30 € in treatment *PE0* (12.35 € for actors, and 9.90 € for bystanders), 14.79 € in treatment *PE20* (14.76 € for actors, and 14.84 € for bystanders), 13 € in treatment *NE20* (13.29 € for actors, and 12.61 € for bystanders) and 20.71 € in treatment *NE60* (18.32 € for actors, and 23.90 € for bystanders).

no treatment differences. The actors' as well as the bystanders' payoff is the initial endowment $e_A$ and $e_B$, respectively.

In recent years considerable experimental evidence has been collected on subjects' provision behavior in public goods games, showing systematic deviations from this prediction. In one-stage games subjects typically contribute about 40-60% of their endowment. In repeated interactions subjects typically start off in about the same range, but over the course of the interaction contributions decrease to very low levels (Fehr and Gächter 2000; Keser and van Winden 2000; Brandts and Schram 2001; Fischbacher, Gächter et al. 2001; Zelmer 2003; Chaudhuri 2011). The obvious explanation that the low contribution levels in advanced periods are due to subjects' learning of the free-riding incentives has lost bite in the observation of a considerable increase of cooperation after a "restart" (Andreoni 1988). A prominent explanation of contribution patterns is *conditional cooperation*. Conditionally cooperative subjects cooperate if they expect other subjects to cooperate as well and free-ride otherwise. In direct tests, about half of the subjects have been classified as conditional cooperators, while 20-33% were identified as free-riders (Fischbacher, Gächter et al. 2001; Kurzban and Houser 2005; Fischbacher and Gächter 2010).

Although conditional cooperation is a well established explanation for the contribution patterns observed in public goods games, there is no general theory of conditional cooperation. The modifications of the standard model that allow for (conditional) cooperation in equilibrium are based on different forces driving this behavior. Kreps, Milgrom et al. (1982) showed that when abandoning common knowledge of rationality and assuming incomplete information about the other player's type, cooperation may occur in a sequential equilibrium of the finitely repeated prisoners' dilemma game. Narrowly self-interested players conditionally cooperate in equilibrium if both of them believe that there is a small chance that the opponent achieves extra utility from mutual cooperation (*altruist*) or adopts a tit-for-tat strategy. In fact, altruistic players do not have to exist; the belief of their existence is sufficient to sustain cooperation. If altruistic players actually exist, defection may no longer be a best reply to cooperation and conditional cooperation may even occur in the one-stage prisoners' dilemma game (Andreoni and Miller 1993; Cooper, DeJong et al. 1996). This explanation exclusively hinges on the cooperation preferences of active players and the information thereof. Adding passive players should not affect the results. If this was the (sole) driving force of conditional cooperation, we should therefore not see treatment differences.

In recent years various models assuming a wider notion of self-interest have been proposed (Sobel 2005; Fehr and Schmidt 2006). Players act to maximize their utility, which is not solely influenced by their monetary gain, but also by concerns for other players (*interdependent* or *other-regarding preferences*) (Rabin 1993; Levine 1998; Fehr and Schmidt 1999; Bolton and Ockenfels 2000; Charness and Rabin 2002; Dufwenberg and Kirchsteiger 2004). Out of different motivations, these theories allow for conditionally cooperative behavior to be an equilibrium of one-stage games, either in terms of "conventional" or psychological game theory (Geanakoplos, Pearce et al. 1989).

Rabin's (1993) model of intention based reciprocity is pioneering as it allows for multiple equilibria in psychological two player games in which players act reciprocally based on the other player's intentions. If a player is perceived to be kind, the opponent wants to be kind too and vice versa. Consequently conditionally cooperative behavior may be observed in equilibrium. Rabin's model has been extended and generalized, e.g. by Charness and Rabin (2002), by Dufwenberg and Kirchsteiger (2004) to N-person extensive form games and by Falk and Fischbacher (2006) to combine inequity aversion with intentions. In any way the perception of the opponent's intentions depends on the player's belief about why the opponent is acting this way. Levine (1998) assumes in his model of interdependent preferences that a player's utility not only depends on her monetary payoff, but also on the other players' types. A player acts more cooperatively towards an altruistic than towards a spiteful type. When another player's type is not known, initial beliefs are updated through observed behavior. What happens – ceteris paribus – when passive bystanders are added to intention based models or Levine's model of interdependent preferences? Although we cannot rule out that actors have initial beliefs about the kindness and the spitefulness of the bystanders, the bystanders' passivity does not allow collecting any information to update these priors. Evidently, bystander effects are not in the core focus of these theories and they do not allow us to predict how treatment differences affect contributions.

Models of inequity aversion assume that players compare their payoff to the payoffs of the other players, either individually (Fehr and Schmidt 1999) or on an aggregated level (Bolton and Ockenfels 2000). In the Fehr and Schmidt (1999) model, actors gain utility from their monetary payoff and disutility both from having a payoff disadvantage and from having a payoff advantage in comparison to each of the other players. Complete free-riding of all actors also constitutes an equilibrium in the Fehr-Schmidt model, but there may be additional equilibria of the one-shot game in which at least some actors contribute a positive amount.

These equilibria require a sufficient number of actors who sufficiently suffer from a payoff difference to their advantage. These actors have no incentive to deviate to a lower contribution because the monetary advantage is more than eaten up by the disutility of outperforming their peers. Then players no longer have the dominant strategy of free-riding, but may cooperate if they expect others to cooperate, and free-ride if they expect others to free-ride. This implies that inequity averse players act as conditional cooperators.

In models of inequity aversion passive bystanders potentially affect actors' behavior, because payoff comparisons are independent of the strategic possibilities of the other players. As we show in more detail in Appendix B, the predictions of the Fehr-Schmidt model vary with the treatment parameterizations. All treatment variations, even the one with unaffected bystanders (*Nox7*), strengthen the requirements for equilibria with positive contributions, as compared to the standard public goods game without bystanders (*Nox4*). This means that adding bystanders reduces the chances to observe conditionally cooperative behavior. There are, however, treatment differences. When bystanders have high initial endowments (*PE40* and *NE60*) or when endowments are equal and the externality is positive (*PE20*) there are no equilibria with positive contributions and complete free-riding constitutes the unique equilibrium. Thus, inequity aversion leaves no room for conditional cooperation. If bystanders are unaffected (*Nox7)*, have low endowments (*PE0*) or are equally endowed and negatively affected (*NE20*), equilibria with positive contribution levels (10 or 20) are possible if actors are sufficiently averse against advantageous inequality. Thus, in the presence of bystanders we should most likely expect to observe conditionally cooperative behavior in these treatments.

Bolton and Ockenfels (2000) propose an alternative model of inequity aversion, which also allows for equilibria with conditionally cooperative behavior. Next to their monetary payoff, subjects are motivated by the comparison of their individual payoff to the average payoff of all players. In that comparison, achieving exactly the average payoff creates the highest utility, whereas achieving a payoff below (above) average creates incentives to reduce (increase) cooperation. Also in the Bolton-Ockenfels model, complete free-riding is an equilibrium. But, if the probability that others cooperate is high enough, cooperation is a best reply. How does the addition of passive bystanders affect the payoff average? In *Nox7, PE0,* and *NE20* actors' payoff is (weakly) above the average payoff of all players (actors and bystanders together). By increasing the contribution an actor – ceteris paribus – increases all other actors' payoffs in each of these three treatments. In *Nox7* bystanders' payoffs are unaffected by the actors' increase, while in *PE0* also bystanders' payoff is increased and in *NE20* bystanders' payoff is decreased. In these treatments, conditionally cooperative behavior may be expected. In *PE20, PE40,* and *NE60* actors' payoff is (weakly) below the average payoff of all players. Thus, actors want to reduce the payoffs of all other players. Since contributing to the public good increases the payoffs of the other actors and in *PE20* and *PE40* additionally the payoffs of the bystanders, actors are not expected to do so. Thus, also under inequity aversion à la Bolton-Ockenfels conditional cooperation should not be expected in *PE20, PE40,* and *NE60.*

The embedding of our experimental setting into a theoretical framework and the analysis of whether and how the predictions of these theories change – ceteris paribus – when adding bystanders to the public goods game can be summarized in the following hypotheses:

**(H1)** We expect to observe **no treatment differences (**i.e. no effect of the presence and the externality on bystanders) and…

a. … **no cooperation**, when all actors are solely motivated by the maximization of their monetary gains and this is common knowledge.

b. … **conditional cooperation**, when actors believe that some other actors gain sufficient utility from cooperating.

**(H2)** We expect to observe **treatment differences (**i.e. an effect of the presence of and the externality on bystanders) and **conditional cooperation** in the stage game if actors hold other-regarding preferences in the form of inequity aversion with the following treatment differences:

a. In the absence of bystanders (Nox4) cooperation is higher than in any of the treatments with bystanders (including the case of unaffected bystanders, Nox7).

b. In the presence of bystanders cooperation is most likely to be observed in Nox7, PE0, and NE20 and not expected to be observed in PE20, PE40 and NE60.

## 5. Results

To test the hypotheses, we analyze our experimental data at two levels. We first check whether contributions differ between treatments. This leads to a first discrimination between theories. In the next step we analyze how individual participants adjust their contributions to information from past play. Special focus will be put on the benchmarks suggested by the theoretical considerations in section 4. A more detailed explanation of our strategies for parametric estimation is given in the Appendix. The reported two-sided non-parametric Mann Whitney u-tests are performed on mean contributions per subject group, i.e. on the independent observations.

### *5.1 Contributions on an Aggregate Level*

In one respect, all our treatments differ from the baseline: active players are aware of the fact that there are three more participants in their group, be they affected by the active players' decisions or not. Figure 1A shows that this difference alone influences active players' behavior. Contributions are significantly higher when actors are alone than when bystanders are present.[3] Actually, the mere presence of bystanders, although being unaffected, suffices to significantly reduce actors' contributions (see Figure1 B).[4] Both results support (**H2a**).

***Result 1:*** *Not being alone reduces contributions.*

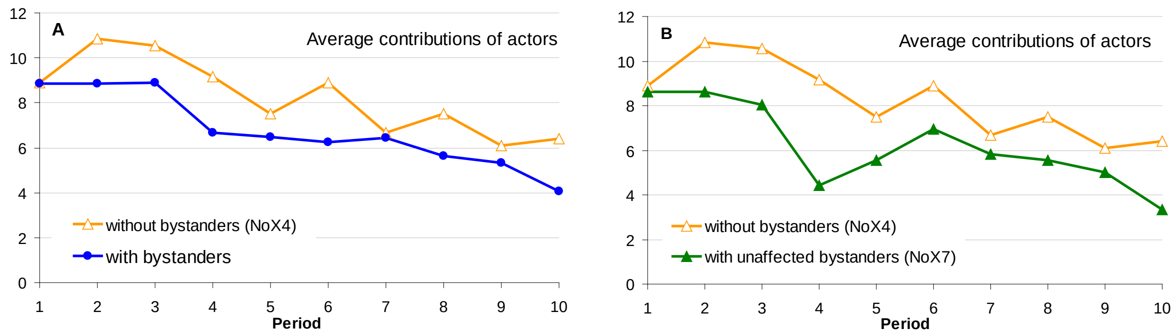


**Figure 1**
**Contributions in the Absence and in the Presence of Bystanders**
A: *Nox4* vs. all other treatments, B: *Nox4* vs. *Nox7*

---

[3] Note that the upward movement from period 1 to 2 in *Nox4* is what one should expect if participants follow the Kreps Wilson logic. In early rounds of a repeated game it pays to give cooperation a chance. An upward movement results if sufficiently many players are positively surprised by the cooperativeness of the remaining group members. Comparable patterns have also been found by other researchers, e.g. Fehr Gächter AER 2000, Figure 3. We find a similar pattern in our treatments *NE20*, *PE40*, *NE60*.

[4] The two-sided Mann Whitney u-test between *Nox4* vs. all other treatments yields p = .0966 (N=63); between *Nox4* vs. *Nox7* it yields p = .0631 (N=18). The regression analyses in Table 2, Models 1 and 2, show that "not being alone" significantly (p<.05) decreases contributions.

| | Pooled | No externality | Positive Externality | | | Negative Externality | |
|---|---|---|---|---|---|---|---|
| | ≠ Nox4 | Nox7 | PE 0 | PE 20 | PE 40 | NE 20 | NE 60 |
| Contribution | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
| not being alone | -.299* | -.425* | .012 | -.426* | -.367* | -.200 | -.369* |
| Period | -.096*** | -.088*** | -.083*** | -.081*** | -.099*** | -.106*** | -.080*** |
| cut 10 | -.596*** | -.820*** | -.814*** | -.836*** | -.929*** | -.943*** | -.794*** |
| cut 20 | .730*** | .421*** | .483** | .542** | .432* | .363* | .484** |
| P model | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 |
| N | 2520 | 720 | 720 | 720 | 720 | 720 | 720 |

**Table 2**
**The Effect of Not Being Alone**[5]
random effects ordered probit, bootstrapped 50 reps, clustered at the group level
Hausman tests on mirror linear random effects models insignificant
reference category *Nox4*
*** p < .001, ** p < .01, * p < .05

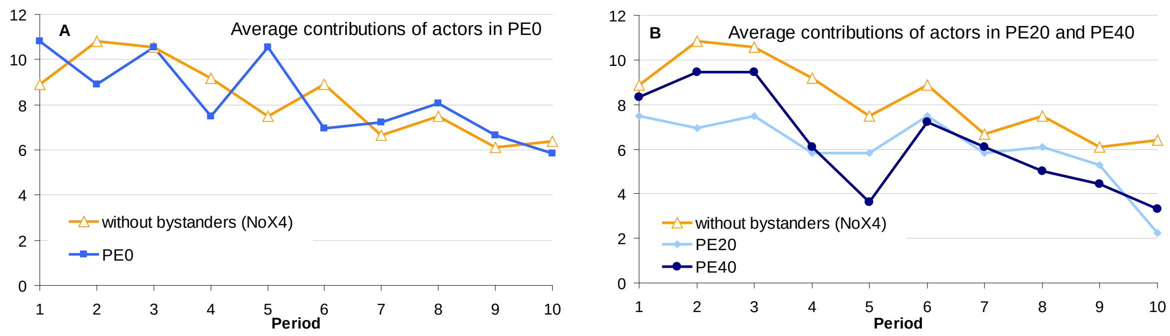


**Figure 2**
**Contributions in the Presence or Absence of Positive Externalities**
A: Comparing *Nox4* with *PE0*, B: Comparing *Nox4* with *PE 20* and *PE40*

If the mere presence of bystanders already influences contributions, what happens if active players' contributions to the public good impact on bystander profit? With positive externalities, contributions are significantly lower than absent any bystanders (*Nox4*) if bystanders are sufficiently rich (in *PE20* and *PE40*); however, if bystanders are "poor" (*PE0*), contributions are not distinguishable from the situation without bystanders (see Figure 2).[6]

[5] A detailed reading and interpretation aid is given in Box B2 in the Appendix.
[6] Mann Whitney u-test, *PE 0* vs. *Nox4*, p = .7235; *PE 20* vs. *Nox4*, p = .0574; *PE 40* vs. *Nox4*, p = .0843 (N = 18 in each comparison). The regression analyses are provided in Table 2, Models 3-5.

Likewise in *NE20* we do not detect any significant difference to the case without bystanders, while we establish a significant difference between *NE60* and *Nox4* (see Figure 3).[7]

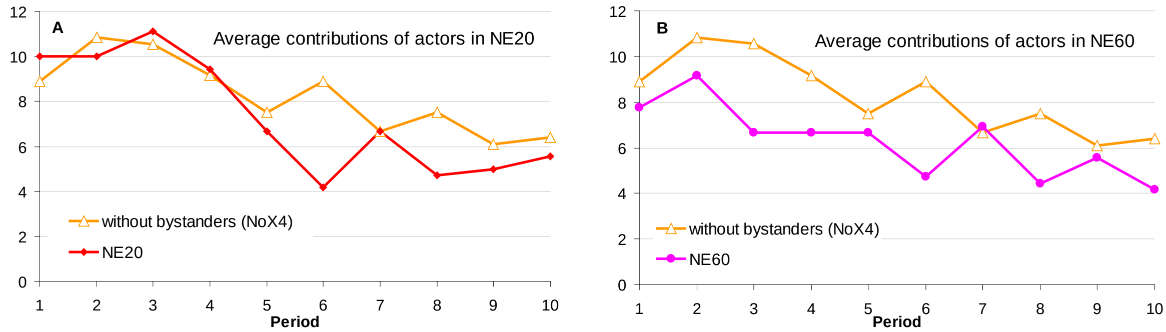


**Figure 3**
**Contributions in the Presence or Absence of Negative Externalities**
A: Comparing *Nox4* with *NE20*, B: Comparing *Nox4* with *NE60*

***Result 2:*** *Contributions display treatment differences, but different from the mere direction of the externality.*

The fact that contributions are strictly positive in all treatments speaks against narrowly self-interested subjects that solely maximize their payoff and believe all other participants to do the same (**H1a**). The fact that contributions differ across treatments speaks against narrowly self-interested subjects who contribute because they believe that others gain utility from cooperation (**H1b**) as the sole explanation. Similarly, our results cannot be simply explained by the direction of the externality. The driving force behind our data seems to be more complex than active players being reticent to impose harm on innocent outsiders, and being encouraged to contribute if this pays a double dividend for outsiders. The analysis of contribution dynamics lets us see driving forces more clearly, and lets us address **(H2)**.

### *5.2 Contributions Dynamics*

*Cooperation Must Pay*

The Kreps, Milgrom et al. (1982) model expects players to cooperate in the finitely repeated game if they believe cooperation will pay. Thus, cooperation should be the more pronounced the more a player experienced that cooperation indeed pays out. By not contributing to the

[7] Mann Whitney u-test, *NE 60* vs. *Nox4*, p = .0698; *NE 20* vs. *Nox4*, p = .5957 (N = 18 in each comparison). The regression analyses are provided Table 2, Models 6&7.

project a participant can make sure to have at least her endowment as a payoff, irrespective of what the remaining active players do. This is why, as a benchmark of whether cooperation has paid, we compare last round's payoff to the endowment of 20.

The fixed effects model in Table 3[8] shows that participants are indeed sensitive to whether cooperation has paid in the past period. The more the payoff exceeded 20, the more a participant increases her contribution in the subsequent period. Actually, given the constant is significant and negative, it is not enough for the payoff to be slightly above 20 to induce participants to increase their contributions. The regression predicts that increases require a payoff of at least 27.[9] While the coefficient for past payoffs in the random effects model is squarely inconsistent, this model indicates that there are no significant treatment differences. Conditional on the difference between the past period's payoff and 20, neither the direction of the externality nor the size of the bystanders' endowment has explanatory power.

| $contr_{i,t} - contr_{i,t-1}$ | random effects | fixed effects |
|---|---|---|
| $\pi_{i,t-1} - 20$ | .139*** | .200*** |
| *Nox7* | -.146 | |
| *PE0* | -.288 | |
| *PE20* | -.143 | |
| *PE40* | -.125 | |
| *NE20* | -.139 | |
| *NE60* | .039 | |
| cons | -.985$^+$ | -1.365*** |
| N | 2268 | 2268 |
| p model | <.001 | <.001 |
| $R^2$ within | .2999 | .1210 |
| $R^2$ between | .0517 | .0450 |
| $R^2$ overall | .2220 | .2227 |

**Table 3**
**Sensitivity to Past Payoff**
Fixed effects instrumental variables regressions
$\pi_{i,t-1} - 20$ instrumented by $\sum_{j \neq i}^{N} contr_{j,t-1}$
standard errors from bootstrap, 50 reps, random draws of entire groups
*** p < .001, ** p < .01, * p < .05, $^+$ p < .1

In line with previous experimental observations we thus find

***Result 3:*** *Subjects conditionally cooperate by increasing contributions when last period's payoff exceeded the endowment and by reducing contributions when contributing has not paid*

8 For a detailed explanation of our estimation strategy see Appendix B3.
9 1.365/.02 = 6.825.

This result clearly supports (**H1b**). However, as mentioned above, if this would already be the entire story, we should not expect to observe the pronounced (unconditional) treatment differences. In the following we thus investigate whether treatment differences may result from payoff comparisons, as suggested by models of inequity aversion.

*Payoff Differences Matter*

If social preferences in the form of inequity aversion contribute to the explanation of contribution behavior, active players should react to the experiences they are making in comparative terms, both in comparison with the other active players and, critically, with passive bystanders. To test this, we analyze the influence of last round payoff differences on the adjustment of participants' contributions over time.

| $contr_{i,t} - contr_{i,t-1}$ | random effects | fixed effects |
|---|---|---|
| $\pi_{i,t-1} - \frac{1}{3}\sum_{j \neq i} \pi_{j,t-1}$ | .186*** | .248*** |
| *Nox7* | -.309 | |
| *PE0* | -.278 | |
| *PE20* | -.309 | |
| *PE40* | -.278 | |
| *NE20* | -.216 | |
| *NE60* | -.123 | |
| cons | -.278 | -.494*** |
| N | 2268 | 2268 |
| p model | <.0234 | <.001 |
| $R^2$ within | .4613 | .2254 |
| $R^2$ between | .0809 | .0734 |
| $R^2$ overall | .3514 | .3521 |

**Table 4**

**Sensitivity to Payoff Comparisons with other Active Players**

Fixed effects instrumental variables regressions

$\pi_{i,t-1} - \frac{1}{3}\sum_{j \neq i} \pi_{j,t-1}$ instrumented by $\sum_{j \neq i}^{N} contr_{j,t-1}$

z-statistic for coefficient from mixed effects regression of lagged differences to remaining active players on their lagged average contributions 26.28, so that instrument is not weak

standard errors from bootstrap, 50 reps, random draws of entire groups

*** $p < .001$, ** $p < .01$, * $p < .05$, $^{+}$ $p < .1$

In Table 4 the regressor for last period's difference between the own payoff and the mean payoff of the remaining active players is highly significant. The positive coefficient shows adjustment to the mean. If the player has outperformed her peers, she increases her contribution. By contrast, the payoff difference is negative if she had contributed more than the average. The positive coefficient of a negative independent variable implies that such participants reduce their contributions in the subsequent period. The fixed effects regression predicts that participants decrease their contributions as long as the negative constant is not offset by a payoff advantage of at least two over the remaining active players.[10] Again, the (inconsistent) random effects model indicates that, conditional on the comparison with the remaining active players, there are no treatment differences.

***Result 4*: *Payoff comparisons to the other actors guide actors' contributions: actors adjust their contribution in the direction of the mean of other actors' last period's contribution***

| $contr_{i,t} - contr_{i,t-1}$ | random effects | fixed effects |
|---|---|---|
| $\pi_{i,t-1} - \pi_{b,t-1}$ | .127** | .181*** |
| *PE0* | -1.799* | |
| *PE20* | .662 | |
| *PE40* | 3.244* | |
| *NE20* | -.752 | |
| *NE60* | 4.614** | |
| cons | -1.084* | .095$^{+}$ |
| N | 1944 | 1944 |
| p model | .1738 | .0001 |
| $R^2$ within | .2670 | .1103 |
| $R^2$ between | .0547 | .0001 |
| $R^2$ overall | .1965 | .0267 |

**Table 5**
**Sensitivity to Payoff Comparisons with Bystanders**
Fixed effects instrumental variables regressions
$\pi_{i,t-1} - \pi_{b,t-1}$ instrumented by $\sum_{j\neq i}^{N} contr_{j,t-1}$
z-statistic for coefficient from mixed effects regression of lagged differences to remaining active players on their lagged average contributions 40.91, so that instrument is not weak
standard errors from bootstrap, 50 reps, random draws of entire groups
*** p < .001, ** p < .01, * p < .05, $^{+}$ p < .1

The adjustment of actors' contributions in the direction of the mean of other actors' last period's contribution may be driven by subjects holding intention based preferences, interdependent other regarding preferences, by repeated game effects, as discussed above, or by subjects holding social preferences in the form of inequity aversion. Payoff comparison to the bystanders should only matter if the last explanation holds true. In Table 5, we analyze how participants react to past differences between their own and bystander profit. The regressor is significant and positive. Thus, whenever there are bystanders, comparing their payoff with them guides active players' behavior. The direction of the bystanders' influence

[10] .494/.248 = 1.99.

depends on the sign of the payoff difference between actors and bystanders. When this difference is positive, i.e. when actors outperform bystanders, payoff comparisons to bystanders induce active players to increase contributions. When, however, bystanders have outperformed actors, their presence reduces contributions, no matter whether actors' contributions actually affect bystanders or not. Thus, bystander comparisons and in particular the desire not to fall back behind bystanders explain the high contributions in *PE0* and *NE20*, as seen in Figure 2A and Figure 3A, and the lower contributions in *PE40* and *NE60*, as seen in Figure 2B and Figure 3B. Both observations are supportive for (**H2b**).

The (inconsistent) random effects model in Table 5 suggests that active players are not only sensitive to comparative performance in the immediate past, but also react to the comparative position to the bystanders. By the design of the experiment in *PE40* and *NE60* active players are always outperformed by bystanders, hence $\pi_{i,t-1} - \pi_{b,t-1}$ is always negative. The positive treatment effect indicates that the reaction to payoff comparisons with bystanders is dampened. Active players do not reduce their contributions whenever they have been outperformed by bystanders, but only if they have fared much worse than bystanders.[11] By the same token, if the design of the experiment already makes sure that active players outperform bystanders (i.e. in *PE0*), they do not always react to this experience by an increase in contributions. This difference must be substantial if it is to trigger an upward adjustment.[12]

Figure 4 shows in which way payoff comparisons with passive bystanders are critical for treatment differences. In two treatments, contributions are not significantly different from the baseline *Nox4*: in *PE0* and in *NE20*. This is precisely the treatments where the active players (almost) always outperform bystanders. In the remaining treatments, at least some active players sometimes are worse off than bystanders. These are the treatments where contributions are lower than in a group with no bystanders.

***Result 5:*** *Payoff comparisons to bystanders guide actors' contributions. In particular, actors do not want to fall behind bystanders*

---

[11] In *PE40*, this requires a difference of more than 25 tokens (3.244/.127 = 25.54). In *NE60*, the model predicts a reduction of contributions if bystanders outperformed the active player by at least 36 tokens (4.614/.127 = 36.33). In both treatments, the cutoff is below the mean difference between active players and bystanders.

[12] Specifically, the difference must be above 14 tokens (1.799/.127 = 14.13). This is, however, below the mean difference.

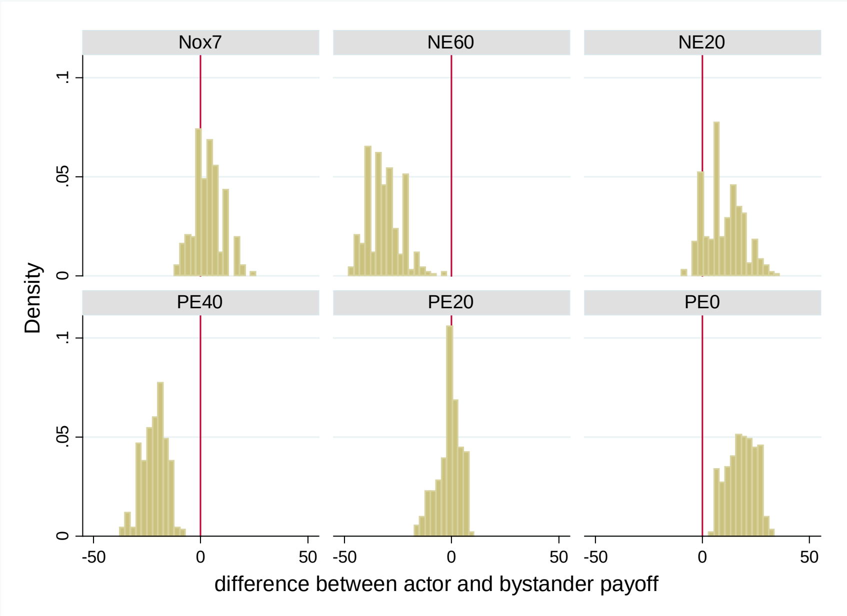


**Figure 4**
**Difference Between Actor and Bystander Payoff**
Values to the left of 0 result from actors having a lower payoff than bystanders
Values to the right of 0 result from actors having a higher payoff than bystanders
histograms are over individual differences per active player and period

## 6. Conclusions

Local public goods frequently spill over to bystanders by either bestowing a windfall profit or by inflicting harm on them. In this paper we show that externalities on bystanders reduce the willingness to contribute to a public good whenever actors risk a competitive disadvantage, compared with bystanders. Actually, the mere salience of a benchmark group suffices to dampen the willingness to contribute to a public good. We explain these findings by the interaction of two effects: sensitivity to the cooperativeness of other active players and, at the same time, the concern for comparative performance. Individual payoff comparisons with respect to the other actors as well as individual bystanders drive contributions down. As long as they cannot fall behind bystanders, actors contribute as much to their joint project as if bystanders were not present. Even a windfall profit for bystanders may be tolerable as long as the additional gain does not make bystanders more prosperous than those contributing to the public good. By contrast if active players run the serious risk of receiving a smaller payoff than bystanders, this second conflict induces them to contribute less.

One must, of course, be cautious when deriving recommendations for institutional design from lab results; context may well suggest otherwise. That said, our findings might help policy makers understand why some social dilemmas are particularly hard to dissolve. According to our results, the proper definition of groups is crucial to whom individuals compare themselves. In particular the comparison of *Nox4* with *Nox7* suggests that the perception of a larger reference group matters even if membership is not at stake. It suffices if a political intervention changes which reference group is made salient. If the reference group transcends the set of active members, voluntary contributions to a public good entail two risks at a time: the risk of being exploited by free-riders at the interior, and the risk of falling behind the external benchmark. Note that such interventions leave the structure of the game unchanged. All that is affected is group construction. When a village close to the national border joins a cross-border association of municipalities, as in the Trans European regions, other members of the association become a natural benchmark for comparisons. A striking example was German reunification. Economically, the new Länder would have been much better off had they tolerated a considerably weaker currency, substantially lower wages, and a less generous system of social security. Yet politically it was impossible not to treat them equally. At the then time the issue was not discussed in these terms. Yet arguably the willingness to voluntarily contribute to the many public goods without which an industrialized country cannot thrive could not have been created in the East, had Easterners had reason to perceive themselves as citizens of second-class.

By contrast, if the difference in prosperity is unquestionable, the fact that they are bestowing a windfall profit on outsiders might not prevent groups from contributing to a joint project. This might, for instance, explain why most industrialized countries have helped create international organizations like the World Trade Organization, even if their efforts also were to the benefit of threshold countries. While the willingness to tolerate positive spillovers may often be normatively desirable, a lack of reticence to impose harm on outsiders is more troubling. Actually, our experimental findings even indicate that insiders might find the negative side effect desirable, precisely because this helps reduce the original distance between themselves and another group. Against this backdrop, it is at least partly comforting that the effect seems less likely if insiders perceive outsiders as inferior. If their individual and social superiority is not at risk, they are not likely to cooperate more fiercely precisely because this harms outsiders.

## Appendix A Instructions

The following instructions are a translation of the German instructions to *PE20*

**Instructions**

**General Information**

- At the beginning of the experiment you will be randomly split into **3 groups of 7 members**. During the whole experiment you will only interact with members of your group.
- The experiment consists of **3 phases**. First you will be informed about phase 1. You will learn about the rules of the next phase as soon as the previous phase has been terminated. **Please note**: The decisions you make in one phase **do not affect** the range of possibilities you have at your disposal in any later phases.

**Information for phase 1:**

- There are **two types** of players: **active** and **passive players**. There are **4 active players** and **3 passive players**. At the beginning of phase 1 it will be randomly determined whether you are an active or a passive player. Your type will remain unchanged for the whole duration of phase 1.
- You play **10 rounds**, every round will have the same structure.
- Each **active** and each **passive** player receives an **endowment of 20 points in each round.**

**Active players**: Each active player has to decide how many of the 20 points he/she wants to **contribute** to the public good. All points contributed to the public good will be **multiplied by 1.6** and **equally** split among all 4 active players, i.e. for every point contributed to the public good by an active player, every **active player receives 0.4 (=1.6/4).** Points not contributed to the public good will stay with the player. More precisely, each active player has to choose one of the following three options:

- **Contribute** 0 points and keep 20 points,
- **Contribute** 10 points and keep 10 points or
- **Contribute** 20 points and keep 0 points

**Passive players**: Passive players cannot contribute to the public good. The payoff of the passive players depends on the contributions of the active players. For each point contributed to the public good by an active player, each passive player receives **0.2 points**.

**Payoff per round:**

| | |
|---|---|
| for **active players:** | **20 –** points contributed **+ 0.4** x sum of the contribution of all active players |
| for **passive players:** | **20 + 0.2** x sum of the contribution of all active players |

**Example**

If the four active players contribute 0, 10, 10 und 20 (arranged by amount), the sum of contributions by all active players is 40 and each active player receives 0.4x40=16 from the joint project. The individual payoffs per round of the **active players** depend on the amounts contributed and are:

- for the player who contributed 0: 20 – 0 + 16 = 36
- for the player who contributed 10: 20 – 10 + 16 = 26 and
- for the player who contributed 20: 20 – 20 + 16 = 16.
-

The payoff per round for each **passive player** is 20 + 0.2x40 = 28.

**Payoff**

Each player receives a base rate of **€ 4** once. At the end of the experiment the points will be paid in Euro with the exchange rate: **10 points** are **0.15 €**.

All other *PE* instructions were analogous. Instructions in *NE* differ in the passage describing the passive player, the passage describing the payoff and the example. In *NE20*, these parts read as:

**<u>Passive players</u>**: Passive players cannot contribute to the public good. The payoff of the passive players depends on the contributions of the active players. For each point contributed to the public good by an active player, <u>each</u> passive player receives a deduction of **0.2 points**.

**Payoff per round:**

| | |
|---|---|
| for **active players:** | **20** – points contributed **+ 0.4** x sum of the contribution of all active players |
| for **passive players:** | **20** – **0.2** x sum of the contribution of all active players |

**<u>Example</u>**

If the four active players contribute 0, 10, 10 und 20 (arranged by amount), the sum of contributions by all active players is 40 and each active player receives 0.4x40=16 from the joint project. The individual payoffs per round of the **active players** depend on the amounts contributed and are:

- for the player who contributed 0: 20 – 0 + 16 = 36
- for the player who contributed 10: 20 – 10 + 16 = 26 and
- for the player who contributed 20: 20 – 20 + 16 = 16.

The payoff per round for each **passive player** is 20 – 0.2x40 = 12.

Instructions in the control treatment *Nox7* differ in the passage describing the passive player, the passage describing the payoff and the example. These parts read as:

**<u>Passive players</u>**: Passive players cannot contribute to the public good. The payoff of the passive players does also not depend on the contributions of the active players.

**Payoff per round:**

| | |
|---|---|
| for **active players:** | **20** – points contributed **+ 0.4** x sum of the contribution of all active players |
| for **passive players:** | **20** |

**<u>Example</u>**

If the four active players contribute 0, 10, 10 und 20 (arranged by amount), the sum of contributions by all active players is 40 and each active player receives 0.4x40=16 from the joint project. The individual payoffs per round of the **active players** depend on the amounts contributed and are:

- for the player who contributed 0: 20 – 0 + 16 = 36
- for the player who contributed 10: 20 – 10 + 16 = 26 and
- for the player who contributed 20: 20 – 20 + 16 = 16.

The payoff per round for each **passive player** is 20

## Appendix B. Supplementary material

### Predictions of the Fehr-Schmidt (1999) model

In the presence of bystanders, inequity may occur with respect to both the other actors and the bystanders. Then, the Fehr-Schmidt model reads as:

$$
\begin{aligned}
u_i^A &= \pi_i^A \\
&\quad - \frac{\alpha_i}{n_A + n_B - 1}\left[\sum_{\substack{j=1 \\ j \neq i}}^{n_A} \max\{\pi_j^A - \pi_i^A, 0\} + n_B \max\{\pi^B - \pi_i^A, 0\}\right] \\
&\quad - \frac{\beta_i}{n_A + n_B - 1}\left[\sum_{\substack{j=1 \\ j \neq i}}^{n_A} \max\{\pi_i^A - \pi_j^A, 0\} + n_B \max\{\pi_i^A - \pi^B, 0\}\right] \\
&\qquad \beta_i \leq \alpha_i, \ \ 0 \leq \beta_i < 1
\end{aligned} \tag{3}
$$

The utility of actor $i$ is composed of the actor's monetary payoff $\pi_i^A$, reduced by the utility loss from disadvantageous payoff differences (second line in (3)) and the utility loss from advantageous payoff differences (third line in (3)). Actors weight disadvantageous inequality with $\alpha_i$ and advantageous inequality with $\beta_i$ .

The unique Nash equilibrium under the assumption of monetary payoff maximization, i.e. complete free-riding of all actors, is also an equilibrium in this version of the Fehr-Schmidt model. In their Proposition 4 Fehr and Schmidt (1999) derive necessary conditions for equilibria in the standard public goods game. We extend this proposition to the case of bystanders and apply it to our treatment parameterizations. As in FS, let *k* denote the number of players with $a + \beta_i < 1$. These players have no (marginal) incentive to contribute to the public good, because the cost of 1 of contributing one unit is not compensated by the monetary return $a$ of this unit and the non-monetary benefit from reducing advantageous inequality $\beta_i$. For these players free-riding is a dominant strategy. If the number of these "free-riders" is sufficiently high, there is no cooperation in equilibrium. Precisely, if

$$
k > \begin{cases} \tfrac{1}{2}a(n_A - 1) & \textit{for treatment } Nox4 \\ \tfrac{1}{2}a(n_A - 1) + \tfrac{1}{2}n_B b & \textit{for treatments } Nox7, PE0, NE20 \\ \tfrac{1}{2}a(n_A - 1) - \tfrac{1}{2}n_B b - n_B(1-a) & \textit{for treatments } PE20, PE40, NE60 \end{cases} \tag{4}
$$

there is a unique equilibrium with zero contributions for all actors. The critical value of *Nox4* corresponds to the one in the standard public goods game (cf. Prop. 4b of FS). In the presence of bystanders the comparison to bystanders enters the actors' profitability calculations. The number of "free-riders" that is "tolerable" for an equilibrium with positive contributions then depends on the number of bystanders $n_B$ and the difference in the marginal benefits of actors and bystanders $a-b$. The presence of unaffected bystanders ($n_B > 0$ and $b = 0$), as in *Nox7*, does not change the critical value of *k* as compared to *Nox4*. Absent any bystanders ($n_B = 0$), the second and third distinction in inequality (4) coincide with the first one.

Table B1 provides the critical values for *k* in our treatment parameterization. The negative critical values in *PE20*, *PE40*, and *NE60* express that the equilibrium with zero contributions of all players is the unique equilibrium, even if there is no single player satisfying $a + \beta_i < 1$. In *Nox4, Nox7, PE0*, and *NE20,* however, there might be positive contribution equilibria, but only if there are no players satisfying $a + \beta_i < 1$. The final column of table B1 shows that in these treatments equilibria with positive contributions are possible if all actors heavily suffer from advantageous inequality, i.e. have very high $\beta$ values. All treatments demand for even more extreme aversion to advantageous inequality than the control *Nox4*. In their part (c) of proposition 4, FS additionally deduce conditions for the existence of asymmetric equilibria in which the players with $a + \beta_i < 1$ contribute 0, while the others contribute positive amounts. These are possible, if the number of players with $a + \beta_i < 1$ is even lower than the bounds deduced in (4). These calculations can also be generalized to the case of bystanders. We will, however, refrain from presenting them here, because – as shown in table B1 – our treatments do not "tolerate" even a single player with $a + \beta_i < 1$. This means that our treatments do not allow for asymmetric equilibria.

| Treatment | Critical value for the number *k* of players with $a + \beta_i < 1$ | Minimum aversion against advantageous inequity for (10,10,10,10) and (20,20,20,20) to be a FS-equilibria |
|---|---|---|
| *Nox4* | 0.6 | $\beta_i > \frac{3}{5} = 0.60 \quad \forall i$ |
| *Nox7* | 0.6 | $\beta_i > \frac{3}{4} = 0.75 \quad \forall i$ |
| *PE0* | 0.9 | $\beta_i > \frac{2}{3} \approx 0.67 \quad \forall i$ |
| *PE20* | -1.5 | - |
| *PE40* | -1.5 | - |
| *NE20* | 0.3 | $\beta_i > \frac{6}{7} \approx 0.86 \quad \forall i$ |
| *NE60* | -0.9 | - |

**Table B1**

**Ranges of inequity parameters**

**Box B.1** Estimation strategy for the parametric analysis of treatment effects

Each participant decides every period how much she wants to contribute. Per subject, data is therefore correlated over time. We capture this relatedness by a subject specific error term, i.e. by a random effects model. Moreover if we analyse contributions, our dependent variable only has 3 expressions: 0, 10, and 20. (Random effects) ordered probit then is the appropriate functional form. We thus work with the following model:

$$y_{it}* = \mathbf{x_{it}}\beta + \alpha_i + \varepsilon_{it}$$

where $y_{it}*$ is a latent variable that varies over participants and periods. $\mathbf{x_{it}}$ is a vector of period and participant specific explanatory variables, with corresponding coefficient vector $\beta$. $\alpha_i$ is a participant specific error term, while $\varepsilon_{it}$ is residual error. Two cutoffs $\gamma_1$ and $\gamma_2$ are estimated with the model. The model predicts variable $y_{it}$ to have expression 0 with probability $p(0) = cdf(\gamma_1 - y_{it}*)$. By the same token, contributions of 20 are predicted with probability $p(20) = 1 - cdf(\gamma_2 - y_{it}*)$, and $p(10 \mid y_{it}*) = 1 - p(0 \mid y_{it}*) - p(20 \mid y_{it}*)$.

Participants stay together in groups of four (seven) over the entire game, which causes our data to be related within groups. In principle, a GLS mixed effects model would be the appropriate way to correct standard errors. Unfortunately, for random effects ordered probit models, there is no generally acknowledged way to do this. We therefore revert to bootstrapping, with drawings at the group level. This gives us standard errors that correct for the relatedness of observations within groups.

This approach still has one limitation. There is no generally acknowledged fixed effects estimator for ordered probit models. Consequently, we are also unable to perform the Hausman test. We must assume that $\alpha_i$ and $\mathbf{x_{it}}$ are uncorrelated. As a double check we run both a random effects and a fixed effects model that ignore the fact that our data only has three expressions and is clustered, and perform the Hausman test on this mirror model.

| **Box B2: Reading and interpretation aid** of Table 2 |
|---|
| As explained in Box B1, the latent variable does not directly map to probabilities, but to a z-standardised normal distribution. To recover the probabilities, one calculates the cumulative distribution function of the lower cutoff, minus the predicted value of the latent variable, to get the probability of a contribution of 0. By the same token, the probability of contributing 20 results from 1 minus the cdf of the upper cut, minus the predicted value of the latent variable. The probability of contributing 10 is 1 minus the two other probabilities.<br>Consider the first period in treatment *Nox4*. In this case the first regressor is 0; only the second regressor matters, and has expression 1. The probability that a participant contributes 0 in this situation is given by cdf (-.596 [cut 10] – (1*-.096 [period])) = 30.85%. The probability that a participant contributes 20 is given by 1 – cdf (.730 [cut 20] – (1*-.096 [period])) = 20.44 %. The probability that the participant contributes 10 is 100 – 24.45 – 26.30 = 48.71 %. By contrast, in the remaining treatments, the probability that a participant contributes 0 in the first period is given by cdf (-.596 [cut 10] – (-.299 [not *Nox4*] + (1*-.096 [period])) = 42.03 %. The probability that a participant contributes 20 is given by 1 – cdf (.730 – (-.299 [not *Nox4*] + (1*-.096 [period])) = 13.03 %, which leaves a probability of 44.94 % that such a participant contributes 10. Hence according to the model it is considerably less likely that a participant contributes 0 in the first period of the baseline, and it is considerably more likely that she contributes 20. |

**Box B3:** Estimation strategy for the analysis of contribution dynamics

In the analysis of contribution dynamics, we face a technical challenge. Dynamics express themselves in first differences, i.e. in $contr_t - contr_{t-1}$. We aim at explaining changes in contributions with experiences participants have made in the previous period like $\pi_{i,t-1} - 20$. Now by the design of a public good

$$\pi_{i,t-1} = 20 - contr_{i,t-1} + \mu * contr_{i,t-1} + \mu \sum_{j \neq i}^{N} contr_{j,t-1}$$

Hence if we estimate

$$contr_t - contr_{t-1} = \beta_0 + \beta_1 * \left(\pi_{i,t-1} - 20\right) + \varepsilon_{it}$$

the term $contr_{i,t-1}$ is on both sides of the equation, which generates endogeneity. We could replace this explanatory variable by total contributions of the remaining group members

$$\sum_{j \neq i}^{N} contr_{j,t-1}$$

or, dividing this by 3, by their average contributions, which do not suffer from this problem. Actually if we explain contribution changes by this variable, we have a highly significant result in the expected direction: participants increase their contributions the more, the more the remaining group members have contributed in the previous period. Yet this estimation strategy deprives us of the possibility to discriminate between driving forces, and hence between theoretical explanations for conditional cooperation.

We overcome this problem by instrumentation. All explanatory variables of interest are highly correlated with the average contribution of others to the public good, in the previous period. This already follows from the design of the experiment, and is corroborated if we explain the respective explanatory variable by lagged average contributions of others.[13] We therefore clearly do not have a weak instrument. On the other hand the dependent variable, i.e. the first differences of contributions of participant $i$ in period $t$, are not correlated with the contribution choices of the remaining participants in the previous period. Hence we have a valid instrument.

On top, we face the usual challenges resulting from the fact that the data from public goods is nested in individuals, nested in groups. Ideally we would therefore want to estimate a mixed effects model. Yet there is no generally acknowledged mixed effects model that allows for instrumentation. We therefore estimate a random effects instrumental variables regression (to cater for dependence at the level of individuals), which we bootstrap with random draws of entire groups (to cater for the dependence at the level of groups). Finally differences of estimated coefficients between fixed and random effects models are big enough to cast doubt on the consistency of the random effects estimator. Since the fixed effects estimator works with mean differencing, and thereby removes the time-invariant treatment effects, we report both models.

[13] In a mixed effects regression of the lagged difference of profit from 20 on the lagged average contribution of others, with standard errors for contributions nested in individuals nested in groups, i.e. with random effects for individuals and groups, with have a z-value of 54.43, Wald Chi2 (1) = 2962.22. Note that from this model we do not get an F-statistic.

## References


ABBINK, KLAUS (2005). Fair Salaries and the Moral Costs of Corruption. Advances in Cognitive Economics. B. N. Kokinov. Sofia, NBU Press: ***.

ABBINK, KLAUS, JORDI BRANDTS, et al. (2010). "Intergroup Conflict and Intra-Group Punishment in an Experimental Contest Game." American Economic Review **100**: 420-447.

ANDERSEN, STEFFEN, ERWIN BULTE, et al. (2008). "Do Women Supply More Public Goods than Men? Preliminary Experimental Evidence from Matrilineal and Patriarchal Societies." American Economic Review: Papers and Proceedings **98**: 376-381.

ANDERSON, LISA R., JENNIFER M. MELLOR, et al. (2004). "Social Capital and Contributions in a Public Goods Experiment." American Economic Review: Papers and Proceedings **94**: 373-376.

ANDREONI, JAMES (1988). "Why Free Ride? Strategies and Learning in Public Goods Experiments." Journal of Public Economics **37**: 291-304.

ANDREONI, JAMES and JOHN MILLER (1993). "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma. Experimental Evidence." Economic Journal **103**(418): 570-585.

BLACKWELL, CALVIN and MICHAEL MCKEE (2003). "Only for my own Neighbourhood? Preferences and Voluntary Provision of Local and Global Public Goods." Journal of Economic Behavior & Organization **52**: 115-131.

BOLTON, GARY E. and AXEL OCKENFELS (2000). "ERC: A Theory of Equity, Reciprocity and Competition." American Economic Review **90**: 166-193.

BOLTON, GARY E. and AXEL OCKENFELS (2010). "Betrayal Aversion. Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. Comment." American Economic Review **100**: 628-633.

BORNSTEIN, GARY (2003). "Intergroup conflict: Individual, Group, and Collective Interests." Personality and Social Psychology Review **7**: 129-145.

BRANDTS, JORDI and ARTHUR SCHRAM (2001). "Cooperation and Noise in Public Goods Experiments. Applying the Contribution Function Approach." Journal of Public Economics **79**: 399-427.

BUCHAN, NANCY R., GIANLUCA GRIMALDA, et al. (2009). "Globalization and Human Cooperation." Proceedings of the National Academy of Sciences **106**(11): 4138-4142.

CHARNESS, GARY and MATTHEW RABIN (2002). "Understanding Social Preferences with Simple Tests." Quarterly Journal of Economics **117**: 817-869.

CHAUDHURI, ANANISH (2011). "Sustaining Cooperation in Laboratory Public Goods Experiments. A Selective Survey of the Literature." Experimental Economics **14**: 47-83.

COOPER, RUSSELL, DOUGLAS V. DEJONG, et al. (1996). "Cooperation without Reputation. Experimental Evidence from Prisoner's Dilemma Games." Games and Economic Behavior **12**: 187-218.

CROSON, RACHEL T.A., ENRIQUE FATAS, et al. (2005). "Reciprocity, Matching and Conditional Cooperation in Two Public Goods Games." Economics Letters **87**: 95-101.

DUFWENBERG, MARTIN and GEORG KIRCHSTEIGER (2004). "A Theory of Sequential Reciprocity." Games and Economic Behavior **47**: 268-298.

ELLMAN, MATTHEW and PAUL PEZANIS-CHRISTOU (2010). "Organisational Structure, Communication and Group Ethics." American Economic Review **100**: ***.

ENGEL, CHRISTOPH (2007). "How Much Collusion? A Meta-Analysis on Oligopoly Experiments." Journal of Competition Law and Economics **3**: 491-549.

ENGELMANN, DIRK and MARTIN STROBEL (2004). "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments." American Economic Review **94**: 857-869.
FALK, ARMIN and URS FISCHBACHER (2006). "A Theory of Reciprocity." Games and Economic Behavior **54**: 293-315.
FEHR, ERNST and SIMON GÄCHTER (2000). "Cooperation and Punishment in Public Goods Experiments." American Economic Review **90**: 980-994.
FEHR, ERNST and SIMON GÄCHTER (2000). "Fairness and Retaliation. The Economics of Reciprocity." Journal of Economic Perspectives **14**: 159-181.
FEHR, ERNST and KLAUS M. SCHMIDT (1999). "A Theory of Fairness, Competition, and Cooperation." Quarterly Journal of Economics **114**: 817-868.
FEHR, ERNST and KLAUS M. SCHMIDT (2006). "The Economics of Fairness, Reciprocity and Altruism. Experimental Evidence and New Theories." Handbook on the Economics of Giving, Reciprocity and Altruism **1**: 615-691.
FISCHBACHER, URS (2007). "z-Tree. Zurich Toolbox for Ready-made Economic Experiments." Experimental Economics **10**: 171-178.
FISCHBACHER, URS and SIMON GÄCHTER (2010). "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Good Experiments." American Economic Review **100**: 541-556.
FISCHBACHER, URS, SIMON GÄCHTER, et al. (2001). "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." Economics Letters **71**: 397-404.
FREY, BRUNO and STEPHAN MEIER (2004). "Social Comparisons and Pro-social Behavior: Testing “Conditional Cooperation” in a Field Experiment." American Economic Review **94**: 1717-1722.
GEANAKOPLOS, JOHN, DAVID W. PEARCE, et al. (1989). "Psychological Games and Sequential Rationality." Games and Economic Behavior **1**(1): 60-79.
GREINER, BEN (2004). An Online Recruiting System for Economic Experiments. Forschung und wissenschaftliches Rechnen 2003. K. Kremer and V. Macho. Göttingen**:** 79-93.
GÜTH, WERNER and ERIC VAN DAMME (1998). "Information, Strategic Behavior, and Fairness in Ultimatum Bargaining. An Experimental Study." Journal of Mathematical Psychology **42**: 227-247.
KESER, CLAUDIA and FRANS VAN WINDEN (2000). "Conditional Cooperation and Voluntary Contributions to Public Goods." Scandinavian Journal of Economics **102**: 23-39.
KOSFELD, MICHAEL, AKIRA OKADA, et al. (2009). "Institution Formation in Public Goods Games." American Economic Review **99**: 1335-1355.
KREPS, DAVID M., PAUL R. MILGROM, et al. (1982). "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." Journal of Economic Theory **27**: 245-252.
KURZBAN, ROBERT and DANIEL HOUSER (2005). "Experiments Investigating Cooperative Types in Humans. A Complement to Evolutionary Theory and Simulations." Proceedings of the National Academy of Sciences of the United States of America **102**: 1803-1807.
LEDYARD, JOHN O. (1995). Public Goods. A Survey of Experimental Research. The Handbook of Experimental Economics. J. H. Kagel and A. E. Roth. Princeton, NJ, Princeton University Press**:** 111-194.
LEVINE, DAVID K. (1998). "Modeling Altruism and Spitefulness in Experiments." Review of Economic Dynamics **1**: 593-622.
MASCLET, DAVID, CHARLES NOUSSAIR, et al. (2003). "Monetary and Non-Monetary Punishment in the Voluntary Contributions Mechanism." American Economic Review **93**: 366-380.

OKADA, AKIRA and ARNO RIEDL (2005). "Inefficiency and Social Exclusion in a Coalition Formation Game. Experimental Evidence." Games and Economic Behavior **50**: 278-311.

OSTROM, ELINOR, THOMAS DIETZ, et al., Eds. (2002). The Drama of the Commons. Washington, National Academy Press.

PAGE, TALBOT, LOUIS PUTTERMAN, et al. (2005). "Voluntary Association in Public Goods Experiments. Reciprocity, Mimicry and Efficiency." Economic Journal **115**: 1032-1053.

POTTERS, JAN, MARTIN SEFTON, et al. (2005). "After You - Endogenous Sequencing in Voluntary Contribution Games." Journal of Public Economics **89**: 1399-1419.

RABIN, MATTHEW (1993). "Incorporating Fairness into Game Theory and Economics." American Economic Review **83**: 1281-1302.

SOBEL, JOEL (2005). "Interdependent Preferences and Reciprocity." Journal of Economic Literature **43**: 392-436.

VAN DIJK, FRANS, JOEP SONNEMANS, et al. (2002). "Social Ties in a Public Good Experiment." Journal of Public Economics **85**: 275-299.

WIT, ARJAAN P. and NORBERT L. KERR (2002). ""Me Versus Just Us Versus Us All". Categorization and Cooperation in Nested Social Dilemmas." Journal of Personality and Social Psychology **83**: 616-637.

ZELMER, JENNIFER (2003). "Linear Public Goods. A Meta-Analysis." Experimental Economics **6**: 299-310.