Can we manage first impressions in cooperation problems?

An experimental study on
"Broken (and Fixed) Windows"

Christoph Engel
Sebastian Kube
Michael Kurschilgen

MAX PLANCK SOCIETY

Can we manage first impressions in cooperation problems?

An experimental study on "Broken (and Fixed) Windows"

Christoph Engel / Sebastian Kube / Michael Kurschilgen

March 2011

Can we manage first impressions in cooperation problems?[*]
An experimental study on "Broken (and Fixed) Windows"

Christoph Engel   **Sebastian Kube**   **Michael Kurschilgen**

Abstract

Cooperation problems are at the heart of many everyday situations. In this paper, we propose a very simple and light-handed mechanism to sustain cooperation and test its performance in a rich laboratory environment. The mechanism moderates cooperation by controlling experiences, more specifically, it "manipulates" subjects' initial beliefs by providing them with selective information about (un)cooperative behavior in other, unrelated, groups. We observe that contributions are considerably sensitive to such selective information. First impressions participants happen to make predict subsequent behavior. Our results, however, suggest an asymmetry in the strength of the reaction – which might pose a limit on the effectiveness of the mechanism in natural settings.

JEL: C91, D03, D83, H41, K14, K42

Keywords: Public Good, Behavioral Uncertainty, Conditional Cooperation, Information, First Impressions, Broken Windows

---

1   Introduction

Cooperation problems are at the heart of many everyday situations. For example, when it comes to protecting the environment, defending one's country, generating new knowledge, joining a political party, extending the infrastructure, or exploiting the opposite market side, agents face a social dilemma. Jointly they are best off if everyone contributes her fair share. But individually, free-riding on others' efforts yields the highest payoff. A large number of theoretical and empirical papers have explored how people should be expected to and how they actually do behave in such situations. Many empirical contributions make use of laboratory experiments where subjects participate in a prisoners' dilemma or in a public-good game. Absent institutional interventions like punishment (Fehr and Gächter 2000), contributions are heterogeneous in the beginning, but average cooperation quickly declines and most participants free-ride in the end (e.g. Andreoni 1988). These results have been stress-tested extensively, e.g., with respect to anonymity (Andreoni and Petrie 2004), culture (Herrmann, Thöni et al. 2008), group size (Isaac, Walker et al. 1994), efficiency (Glöckner, Irlenbusch et al. 2010), or framing (Goerg and Walkowitz 2010). Interventions that mitigate the dilemma are not easily designed. Effective interventions tend to be heavy-handed, often altering the incentive structure of the game such that free-riding is not in an individual's self-interest any more (e.g. Ostrom, Walker et al. 1992; Falkinger, Fehr et al. 2000; Fehr and Gächter 2002; Gürerk, Irlenbusch et al. 2006; Glöckner, Irlenbusch et al. 2010). In this paper, we propose a very simple and light-handed mechanism and test its effectiveness by using laboratory experiments.

Our mechanism is based on an observation that has previously been made in public-goods experiments, namely that (initial) group composition and initial cooperation rates significantly affect the future development of cooperation rates in a particular group (Burlando and Guala 2005; Gächter and Thöni 2007; Beckenkamp, Engel et al. 2009). A likely reason for this behavioral pattern is provided by the influential concept of reciprocity and conditional cooperation (Fischbacher, Gächter et al. 2001; Dufwenberg and Kirchsteiger 2004; Falk and Fischbacher 2006; Fischbacher and Gächter 2010, among many others). If persons are sensitive to the behavior of other group members, those groups who cooperate little in the beginning will become even less cooperative over time, while groups with substantial cooperation in the beginning are able to sustain cooperation over time. This suggests that outsiders might moderate cooperation by controlling experiences. Our mechanism is even less invasive in that it is confined to the first impressions subjects happen to make. More precisely, the idea is to "manipulate" initial beliefs by providing participants with selective information about (un)cooperative behavior in other, unrelated, groups. The information affects initial beliefs about others' behavior, which in turn changes how I myself behave initially. A virtuous (or vicious, depending on the kind of information that is provided) cycle starts: via beliefs, our manipulation determines cooperativeness in the beginning, which in turn determines cooperativeness later on.

Our intervention idea is in the spirit of James Q. Wilson's Broken Windows Theory (Wilson and Kelling 1982) which is based on the work of American sociologist Philip Zimbardo. In 1969, Zimbardo abandoned two identical cars in two different locations: the Bronx, NYC and Palo

Alto, California. "The license plates of both cars were removed and the hoods opened to provide the necessary releaser signals" (Zimbardo 1969). In the Bronx, the abandoned car was stripped and demolished after only 26 hours, the result of 23 separate incidents of vandalism. In contrast, the car in Palo Alto still sat unmolested after the course of an entire week. Zimbardo then decided to provide an example of vandalism to the affluent and seemingly non-violent neighborhood of Palo Alto. So he and two graduate students of his took a sledgehammer and started bashing the car. After they had taken the first blow, observers shouted encouragement and finally joined in the vandalism, until the car was completely wrecked (Zimbardo and Ebbesen 1969).

In order to clearly identify the short- and long-run effects of our mechanism we run laboratory experiments, which have the benefit of providing a sufficient degree of control over the environment. The environment that we use has been prominent lately to study social dilemmas and cooperation problems. It is a complex public-good game (more precisely, a voluntary contribution mechanism) with decentralized sanctions (Fehr and Gächter 2000) and counter-punishment opportunities (Nikiforakis 2008). We have chosen this rich environment, since it mimics several potentially important features available in natural environments. Moreover, it is sufficiently complex, so that conflicting behavioral norms might emerge – which allows us to test the effect of both favorable and unfavorable first impressions.

In the baseline, subjects play the game for 10 periods without any exogenous interventions. In two treatments (FAV and UNFAV), prior to making their first contribution decision in the game, we provide subjects with selected data from the baseline. Specifically, we show them graphs of the development of mean contributions over time from selected cooperative (resp. uncooperative) groups. The information that we give is unfavorable for cooperation in treatment UNFAV – similar to Zimbardo, we "attack the car with a hammer" and explore whether "the car will then go to ruins", i.e., whether cooperation breaks down in the public-good experiment. Contrarily, in treatment FAV we reverse the manipulation by exposing subjects to favorable initial impressions – i.e., similar to Wilson, metaphorically speaking we repair the windows in a bad neighborhood, and investigate whether newly arriving inhabitants behave more cooperatively.

We find that cooperation rates strongly react to the treatment manipulation.[1] Unfavorable information substantially reduces initial contributions to the public good. Moreover, the effect is long-lasting and translates into future development of groups' cooperation behavior. Contributions in treatment UNFAV do not recover during the experiment. More importantly even for policy making, the favorable selective information in treatment FAV prevents cooperation rates from decaying over time. Taken together, our results suggest that subjects' behavior is indeed conditional on initial expectations of how others are likely to behave, and that these expectations are open to purposeful intervention.

---

1   As we demonstrate in a second series of experiments (reported below in Section 4), the effect of our manipulation is mediated via subjects' initial beliefs. Beliefs about others' initial contribution decisions are much more pessimistic when unfavorable information is provided.

The next section embeds our study into the existing literature. Subsequently, we explain the design of the experiment. Section 4 presents and discusses the experimental results of our main study, and features an additional experiment that explores whether the effect of our manipulation is mediated by subjects' initial expectations. Section 5 concludes the paper.

2   Related Literature

Our study closely relates to the experimental literature on reciprocity and conditional cooperation. This literature suggests that a substantial fraction of subjects usually consists of conditional cooperators, i.e., people who are willing to cooperate, provided a sufficiently large fraction of the population does the same. (Fischbacher, Gächter et al. 2001; Mengel 2007; Fischbacher and Gächter 2010). For conditional cooperators, information about contribution patterns in other groups has a first- and a second-order effect. Given the uncertainty about the composition of the group of which they happen to be a member, information about other groups helps them form beliefs. Moreover, each participant knows that each other participant has received the same information. This makes it possible to also form a second-order belief, based on knowing that the remaining group members have received the same information. The finding that, in our experiment, cooperation rates are sensitive to the initial information-manipulation is fully reconcilable with the idea that a non-negligible fraction of the population acts as if they were conditional cooperators. This yields further support for the existence and importance of conditional cooperation.

The research question covered here is also closely related to the literature on institutional design, specifically the branch that explores interventions aiming at raising contributions to public goods. Usually, the proposed mechanisms alter the incentive structure such that free-riding becomes less attractive, or even a dominated strategy for self-centered money maximizers. For example, in (Falkinger, Fehr et al. 2000) the payoff structure is changed such that "each individual gets a reward or has to pay a penalty depending on the deviation of its contribution from the mean contribution." In other studies, group composition is changed such that the payoff structure is changed implicitly. For example, in (Gunnthorsdotir, Houser et al. 2007), subjects are re-matched every period according to their cooperativeness in the previous round; which is found to raise cooperation levels. Likewise, if groups have a chance to exclude free-riders, this improves cooperation in a dilemma setting (Cinyabuguma, Page et al. 2005; Croson, Fatas et al. 2008), as does a mechanism that allows members to self-select into groups (Page, Putterman et al. 2005), in particular if free-riders are effectively excluded by a rule that sacrifices a portion of the group income to outsiders (the Red Cross, as it was , Brekke, Hauge et al. 2009). Our study differs from this literature in that we leave the incentive structure of the game completely unchanged. All we alter are the first impressions participants happen to make – which is an option that should naturally be available in any public good game, and in many real-life social dilemmas.

Parts of the legal literature, particularly those at the intersection of law and economics, have also been asking how socially desirable behavior and/or compliance with the law can be brought

about. Again, a prominent approach is to change the incentive structure, e.g., by increasing the expected costs of breaking the law. For example, Braga, Weisburd et al. (1999) report evidence from a field experiment that randomly exposed 12 of 24 matched violent crime places in Jersey City to intense police scrutiny and intervention. In the places chosen, crime rates dropped substantially, while they did not in the unaffected places. In a similar vein, in a series of sociological field experiments, when there were signs of disorder, like graffiti, abandoned shopping carts, littering or bicycles locked where they were not supposed to be, this induced passers-by to also break these and other rules (Cialdini, Reno et al. 1990; Keizer, Lindenberg et al. 2008; Ramos and Torgler 2010). In laboratory experiments, Galbiati and Vertova (2008) demonstrate that cooperation behavior increases when an explicit expectation is spelt out and enforced by (non-deterrent) sanctions (see also Kube and Traxler 2010).

Besides, our findings also underline the power of information, as it has also been observed in other contexts. For example, previous work on voting behavior points out that information in the form of polls (Forsythe, Myerson et al. 1993; Forsythe, Rietz et al. 1996; Klor and Winter 2007) or of cheap-talk electoral campaigns (Corazzini, Kube et al. 2010) affect subsequent voting outcomes. Similarly, information gathered during pre-play communication[2] strongly affects subjects' decision in subsequent coordination games (Blume and Ortmann 2007), as does information from preceding asset-market outcomes (Kogan, Kwasnica et al. 2010) or information about group members' previous decisions (Weber 2006). Interestingly, the literature on coordination games comes to the conclusion that pre-play information promotes Nash-equilibrium play (assuming self-centered money maximizers). By contrast, our experimental results show that the opposite might happen in a cooperation game like the one reported here, where the socially efficient outcome is usually not part of the subgame-perfect Nash equilibrium. Pre-play information may indeed promote socially efficient outcomes.

3   Experimental Design, Procedure, and Behavioral Predictions

Experimental Design: The baseline treatment is a typical public-good game with punishment and counter-punishment opportunities as implemented by Nikiforakis (2008).[3] The basic game features n=4 players and consists of three stages:

At the beginning of stage 1 ("contribution stage"), players are endowed with 20 tokens each. Players then decide simultaneously and independently how much of their endowment they want to contribute to a public account. We denote this decision with c_i. Each token that is contributed

---

2   See alsoCrawford (1998) for a general survey of experiments on communication via cheap talk.

3   We have chosen this design because it provides a rich environment in which different behavioral norms may emerge. Moreover, this design has the advantage to leave sufficient room for testing the effect of both favorable and unfavorable first impressions. Had we tested a mere voluntary contribution mechanism, previous experiments would have indicated that cooperation is very difficult to sustain in the first place. Conversely, had we chosen a public good game with one-step punishment, successful cooperation would have been very likely. By contrast, according to the existing literature, if we also add the counter-punishment option, expectations are in the middle between both extremes.

to the public account increases the payoff of each player in the group by 0.4 tokens (i.e., the MPCR is α=.4). Each token unspent increases a player's own payoff by one token. The preliminary payoff at the end of stage 1 is thus given by $\pi_i^1 = 20 - c_i + 0.4\sum_{h=1}^{4} c_h$.

p_{ij}	0	1	2	3	4	5	6	7	8	9	10
$C(p_{ij})$	0	1	2	4	6	9	12	16	20	25	30

Table 1: Punishment points p_{ij} per player j and associated costs $C(p_{ij})$ for punisher i

At the beginning of stage 2 ("punishment stage"), players are informed about every group member's contribution to the public account. Every player i then has the opportunity to reduce the income of each other group member j by assigning costly punishment points p_{ij}. Each punishment point received reduces a player's income from stage 1 by ten percent. At the same time, each punishment point assigned reduces one's own payoff according to the cost function given in Table 1. The preliminary payoff at the end of stage 2 is thus given by:

$$\pi_i^2 = \pi_i^1 \frac{1}{10}\max\left\{0,\, 10 - \sum_{j \neq i} p_{ji}\right\} - \sum_{j \neq i} C(p_{ij}).$$

At the beginning of stage 3 ("counter-punishment stage"), players observe who punished them by how much in stage 2. They then have the opportunity to counter-punish the punishers by assigning them counter-punishment points cp_{ij}. The punishment technology is the same as in stage 2. Each counter-punishment point received reduces a player's preliminary income from stage 2 by ten percent and each counter-punishment point assigned reduces one's own payoff as given in Table 1, but the cost of counter-punishment also depends on this player's punishment decision.[4] The final payoff at the end of stage 3 is thus given by:

$$\pi_i^3 = \pi_i^2 \frac{1}{10}\max\left\{0,\, 10 - \sum_{j \neq i} cp_{ji}\right\} - \sum_{j \neq i}[C(p_{ij} + cp_{ij}) - C(p_{ij})]$$

In treatments FAV and UNFAV, the underlying game structure remains unchanged. Compared with the baseline the only change is that we append the experimental instructions by one additional page. The page contains selected information about unrelated groups who previously played this game – and players know that the information is actually taken from previous experiments. They are also told that the information is selective and that all players in their group receive the same information. The information we give is the development of mean contributions over time from four selected groups. Furthermore, we provide them with selected data about the number of persons choosing to contribute zero in the first, resp. in the last period, and the average amount of counter-punishment that was meted out in these experiments. The way this data is presented and the selection of the four groups are such that they constitute a positive (i.e. cooperative) impression in treatment FAV, while they constitute a negative (i.e. uncooperative) im-

---

4   That is, if i assigns counter-punishment points to j, the specific costs of the counter-punishment points depend on the number of punishment points that i has assigned to j on stage two. For example, if i had already assigned four points to j on stage 2, assigning him a single counter-punishment point on stage three costs 3 tokens. Hence the costs of counter-punishment are given by $C(p_{ij}+cp_{ij})-C(p_{ij})$.

pression in treatment UNFAV.[5] Thus, the treatment manipulations allow us to study to what extent different contents of pre-play information affects behavior.

Procedure: The experiments were run in two different locations, namely in London (UK) and Bonn (Germany). The baseline treatment was played in both locations. As will be seen below, the cooperation rates in the baseline treatment differ between these two locations. The observed cooperation rates in this environment are rather low in London, while they are rather high in Bonn .[6] Taking advantage of this observed variation between the two locations, we run treatment FAV in London and treatment UNFAV in Bonn. Consequently we refer to the London baseline as BASEFAV and to the Bonn baseline as BASEUNFAV. Apart from the difference in location and thus the language of the instructions, the procedure was as similar as possible in all treatments.

All experiments were programmed using z-Tree (Fischbacher 2007). Subjects in Bonn were invited using ORSEE (Greiner 2004). The composition of subjects with respect to age, gender and field of studies was similar between locations and between treatments. Subjects were not allowed to participate in more than one treatment (between-subject design). When subjects arrived in the lab, they were seated in separate cubicles. The experimental instructions were then handed out to them and read out aloud. This was done to create common knowledge and to ensure that everybody had read and understood the instructions. Additionally, subjects had to take a quiz and could pose comprehension questions in private before the game started. The instructions were written in neutral language, avoiding potentially loaded terms like punishment or public good (cp. Appendix 1).

After subjects had finished reading the instructions, in treatments FAV and UNFAV they were provided with the additional information sheet (cp. Appendix 3 and 4). Subjects were then randomly divided into groups of four and played the above described game repeatedly for ten consecutive periods. A partner protocol was used, i.e., the group composition stayed constant over the entire 10 periods of anonymous interaction. In the end, subjects were privately paid their cumulated earnings and left.

We had 60 participants (3 sessions) in treatment FAV in the Laboratory for Experimental Economics at the University of Bonn in 2009. Participants received their accumulated earnings from the experiment (1 token = Euro 0.04) and an additional show-up fee of 5 Euro. On average, a session lasted 60 minutes and subjects earned Euro 14.46. The four sessions for treatment UNFAV were run in 2010 at the experimental laboratory of Royal Holloway, University of London, UK. We used the same lab and recruited 64 students from the same subject pool as in Nikiforakis (2008). We also used the same software. Again, participants received their accumulated earnings from the experiment (1 token = 0.04 British Pounds). On average, a session lasted 60 minutes

---

5   The Appendix contains the exact wording of these two pages including the graphs, as well as a translation of the general instructions that were used in all treatments.

6   The corresponding data that we use for our baseline treatments BASEFAV and BASEUNFAVare taken from Nikiforakis (2008) and Beckenkamp, Engel et al. (2009), respectively. The data reported from treatments FAV and UNFAV were collected by us. They are novel and not reported elsewhere. We are thankful to Nikos Nikiforakis for providing us with his data.

and subjects earned 9.06 Pounds. In the two baseline treatments, analogous procedures had been used. The datasets comprise 48 subjects in BASEFAV and 68 participants in BASEUNFAV. Additional details on the two baselines can be found in Nikiforakis (2008) and Beckenkamp, Engel et al. (2009), respectively.

Behavioral Predictions: For rational, self-centered money maximizing players, the game at hand is a cooperation problem. Using backwards induction, it is straightforward that the unique subgame perfect equilibrium is to contribute nothing to the public good at the first stage, and not to punish nor counter-punish at subsequent stages. This leads to an equilibrium payoff of 20 tokens per period and player. By contrast, the socially efficient outcome would be achieved if everyone contributed their entire endowment, in which case every player would earn 32 tokens per period. But in that case, each player would have an individual incentive to free-ride on the others' contributions, which would yield him a payoff of 44 tokens, ceteris paribus. Since we announce the number of periods, through unraveling this is also the prediction for the repeated game. The standard prediction of no cooperation not only holds true in the baseline treatment. Providing subjects with additional information about other groups does not change the equilibrium prediction. Therefore, with rational, self-centered money maximizing players we should not expect to observe different behavior in treatments FAV or UNFAV.

By contrast, things might change as soon as we allow for social preferences, in particular if we expect a substantial fraction of the population to act as conditional cooperators. If conditional cooperators are sufficiently optimistic about the cooperativeness of their interaction partners, cooperation gives them higher utility than defection. Facing a public good, conditionally cooperative subjects should therefore base their decision on what they believe other subjects to do. The beliefs are likely to be related to other group members' behavior in the previous period(s). Yet, upon the first encounter (in the first period), a conditional cooperator needs to form initial expectations of how others are likely to behave. This is where our treatment manipulation might make a difference. Pre-play information about other, unrelated groups might influence a participant in forming initial beliefs, which should then guide her first-period decision. If her (conditionally cooperative) group peers were also influenced by the selective pre-play information, then group contributions in the first period are likely to mirror the pre-play information. Her beliefs for the second period will thus be very similar again, and so will contributions be in the second period; and similarly in subsequent periods. In that case, we would expect to observe more cooperation in treatment FAV than in BASEFAV, and less cooperation in treatment UNFAV than in BASEUNFAV throughout the game.

4 Results

In the following, we will first briefly describe the results from the two baseline treatments BASEFAV and BASEUNFAV. As will be seen, observed cooperation rates are low in London while they are high in Bonn. Afterwards, we move on to show that the treatment manipulations have a sub-

stantial and significant influence on subject's behavior. Compared to the respective baseline treatment, cooperation rates are lower in treatment UNFAV while they are higher in treatment FAV. As will be seen, this is likely to be driven by the change in first period's contributions which, in turn, are sensitive to our manipulation of first impressions. In line with this, group's behavior in the first period is strongly correlated with its future development, as we subsequently show. In the last subsection, additional experiments further underline our result by explicitly exploring how beliefs react to the treatment manipulations.

BASELINES:

Figure 1 shows the mean development of contributions over time in BASEFAV and BASEUNFAV. As can be seen, the environment in which subjects interact is complex and allows for different behavioral patterns.[7] Without any manipulation, cooperation rates are rather low in London (BASEFAV). The mean contribution in BASEFAV is 8.72. By contrast, subjects in Bonn (BASEUNFAV) display a high degree of cooperation. They contribute on average 16.45 to the public good. We therefore test in Bonn whether poor first impressions reduce the tendency to contribute much, and thereby trigger a vicious cycle. Conversely we test in London whether promising first impressions increase the willingness to give more initially, and thereby trigger a virtuous cycle.
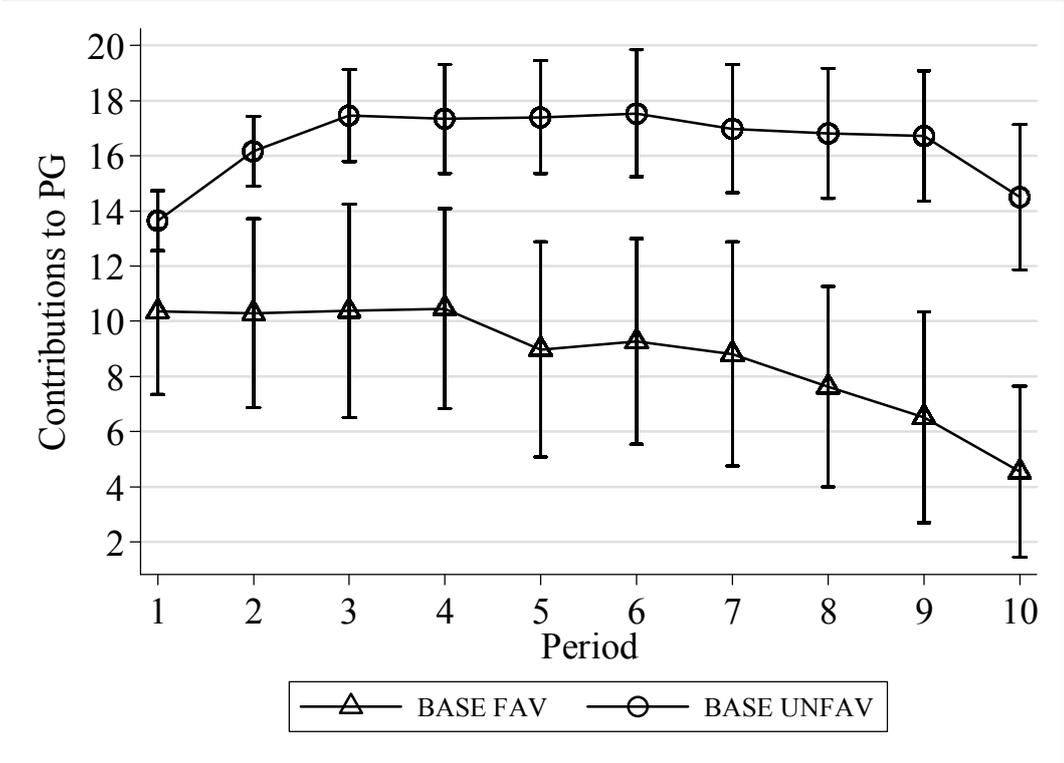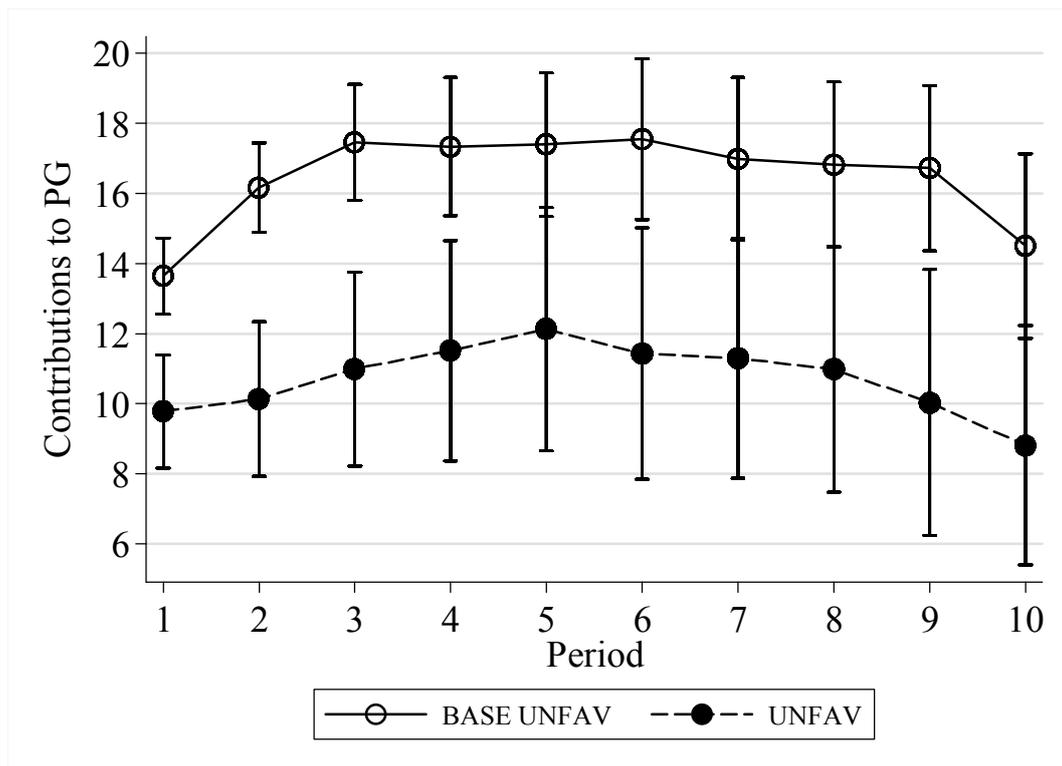

Fig. 1 – Mean Contribution over time in BASEFAV and BASEUNFAV

Notes: With 95% Confidence intervals around the period means. Group clustered standard errors. BASEFAV(Nikiforakis 2008) has 48 subjects (12 groups). BASEUNFAV(Beckenkamp, Engel et al. 2009) has 68 subjects (17 groups).

---

7 While this observation, in particular the behavioral differences between locations, is interesting *per se*, here we do not want to elaborate on this point but would like to refer the interested reader to (Beckenkamp, Engel et al. 2009) for an extensive discussion.

BREAKING WINDOWS:

What happens if we provide subjects with additional information that conveys a negative pre-play impression? As can be seen in Figure 2, this experimental manipulation has a substantial effect on contributions. After "breaking the window" with an "experimental hammer", behavior changes radically. Whereas in BASEUNFAV subjects choose full contribution in 60% (411/680) of all cases, in UNFAV they only do so in 20% (120/600) of all cases. Also at the opposite end of the contribution scale the effect is evident. In BASEUNFAV subjects only in 8.5% (58/680) of all cases contribute less than 5 tokens. With the treatment manipulation, this fraction grows to almost 32% (191/600). Overall, the mean contribution drops significantly from 16.45 in BASEUNFAV down to 10.71 in UNFAV (Mann Whitney ranksum test, N = 32, p=0.0024, two-sided). Hence, our "experimental hammer" indeed creates uncooperative behavior in an otherwise cooperative environment, just like Zimbardo's sledgehammer generated Bronx-like behavior in Palo Alto. Moreover, the effect is long-lasting: the initial manipulation of expectations still leads to considerably lower cooperation even after ten periods of interaction, i.e. after 9 possible opportunities to revise expectations.
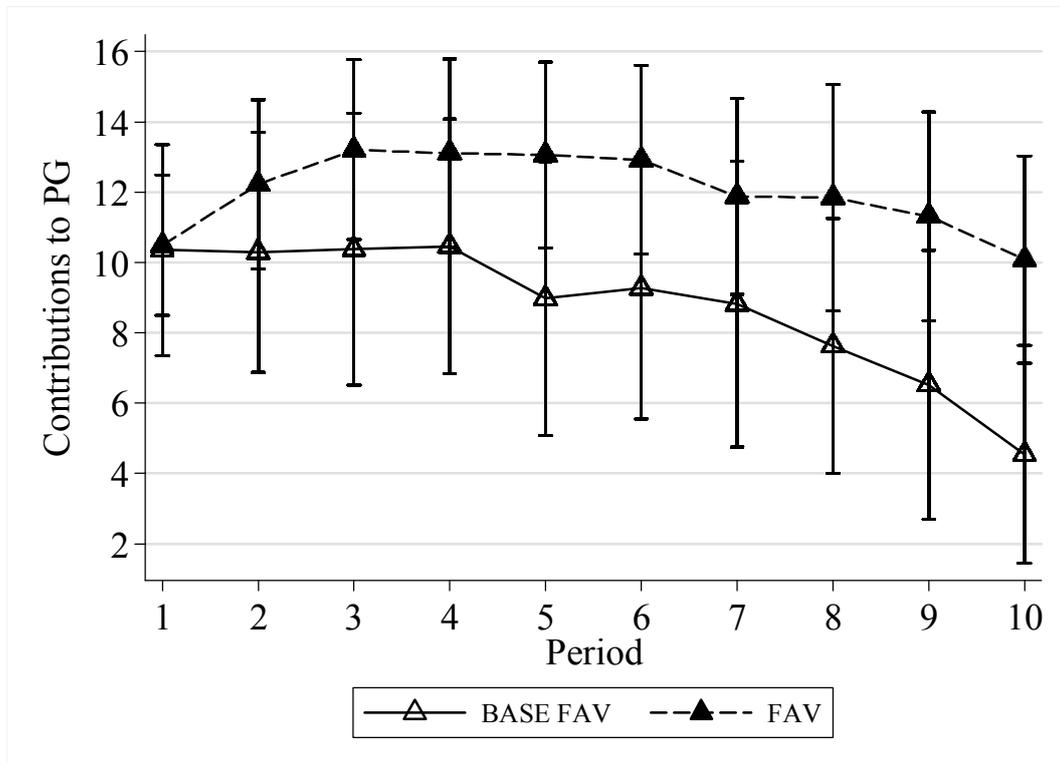

Fig. 2 – Mean Contribution over time in BASEUNFAV and UNFAV

Notes: With 95% Confidence intervals around the period means. Group clustered standard errors. BASEUNFAV (Beckenkamp, Engel et al. 2009) has 68 subjects (17 groups). UNFAV has 60 subjects (15 groups).

FIXING WINDOWS:

From a policy perspective, the opposite situation is even more interesting. Is the mere fact that one "repairs windows" enough to foster cooperation? To test this hypothesis, in treatment FAV we provide participants with favorable first impressions. Inspecting Figure 3, it is plain that repairing windows is less powerful than breaking them. Descriptively, once participants have seen graphs from groups that have cooperated successfully, they contribute more. Average contributions increase from 8.7 in BASEFAV to 12 in FAV. This increase of almost 40% is remarkable, though it falls short of being statistically significant (ranksum test, p=.1567, two-sided). Still, the behavioral pattern seems to be strongly influenced by the favorable manipulation. While cooperation is sustained in FAV, contribution rates quickly decline in BASEFAV. Consequently, from the second period on, mean contributions in BASEFAV are outside the 95% confidence interval of mean contributions in FAV after participants have seen graphs from cooperative groups. Contributions rise sharply for the first three periods. They then stay more or less stable until period six. There is a slight decay for the final periods. Compare this to the data from treatment BASEFAV. Contributions never rise. They decay from the fourth period on. This difference of trends is significant (Mann Whitney over means of first differences, N = 28, p = .0243). Our data thus suggests that fixing windows pays. The beneficial effect only takes longer to unfold.


Fig. 3 – Mean Contribution over time in BASEFAV and FAV

Notes: With 95% Confidence intervals around the period means. Group clustered standard errors. BASEFAV(Nikiforakis 2008) has 48 subjects (12 groups). FAV has 64 subjects (16 groups).

**INITIAL CONTRIBUTIONS AND FUTURE DEVELOPMENT:**

How does this very simple and light-handed mechanism of providing subjects with information from unrelated groups create these strong differences in the overall cooperation rate? If we take a closer look at our data, the answer to this question seems to be twofold.
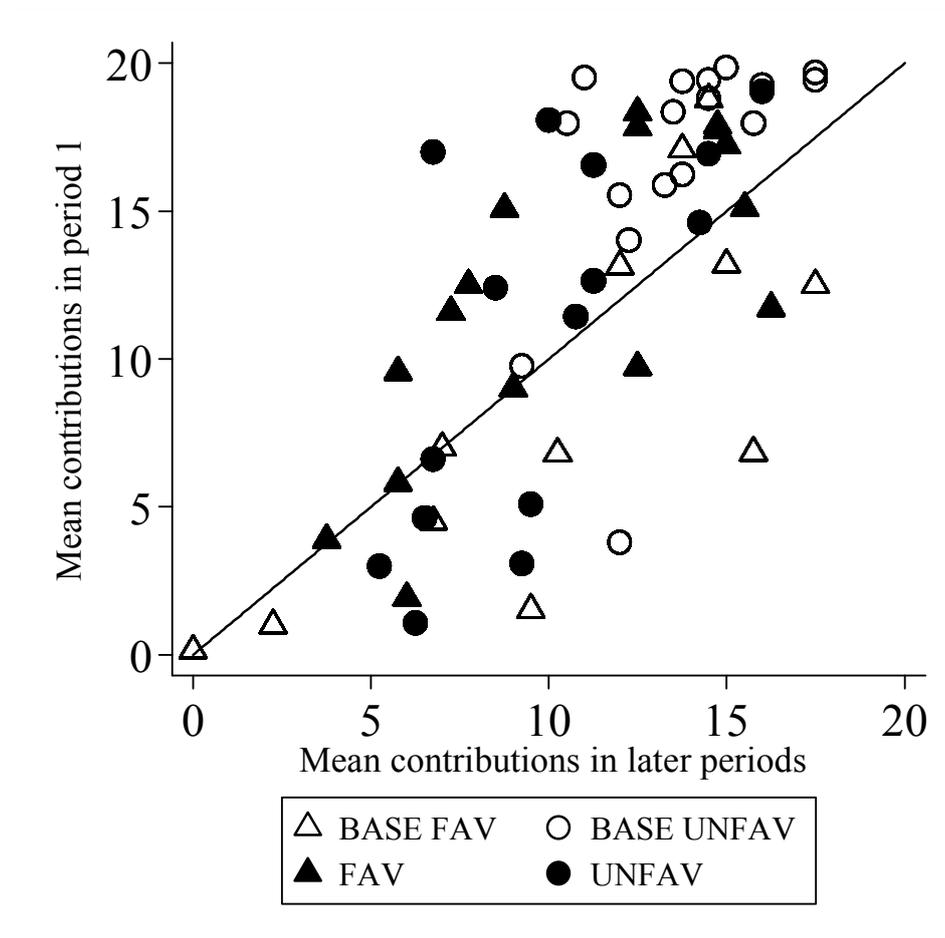
Fig. 4 – Initial contributions vs. future group development

Notes: Every entry in the graph represents one specific 4-person group. There are 12 groups in BaseFav, 17 in BaseUnfav, 16 in Fav and 17 in Unfav. "Later periods" refers to the mean contribution in periods 2 to 10. The 45 degree line depicts all points in which the average contribution in period 1 is identical to the average contribution in later periods.

First, we find that in all treatments contribution behavior in period 1 significantly predetermines cooperativeness in later rounds. This can best be seen when we look at the development in individual groups. Figure 4 compares the average contribution to the public good in period 1 with the average contribution in periods 2 to 10. There is a clear correlation between initial and future cooperation. The correlation is strong and significant, over all treatments (Spearman's rho=.6981, N = 60, p<.001) as well as for each treatment separately.
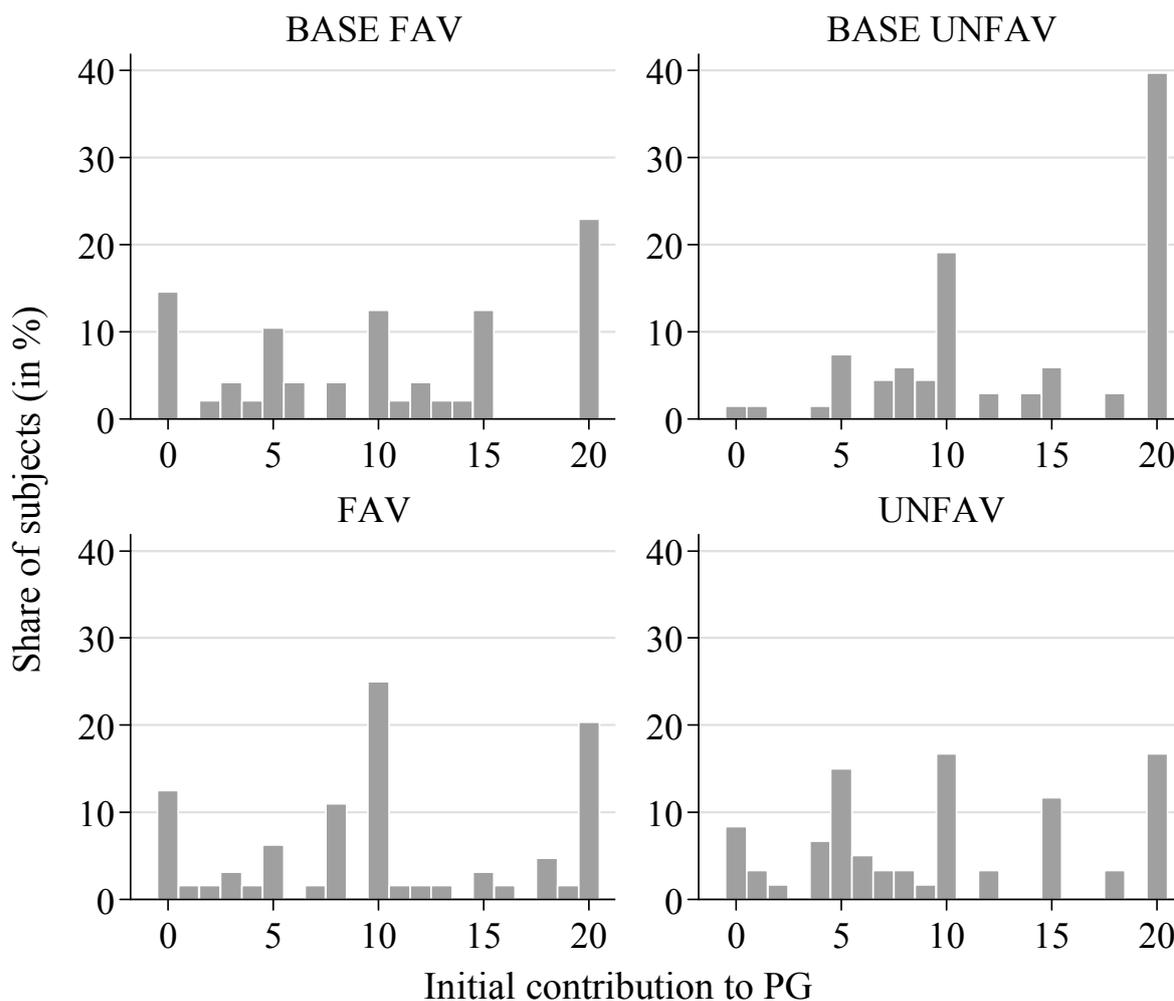
Fig. 5 – Initial contributions to the public good

*Notes:*BASEFAV has 48 subjects, BASEUNFAV 68, FAV 64, and UNFAV 60.

Second, we observe that our experimental manipulation has a substantial effect on initial contributions to the public good (cp. Figure 5). In BASEUNFAV, 40% of subjects go for full contribution in the first period. Yet if they receive the unfavorable information in treatment UNFAV, only 16% do so. The mean contribution in the first period drops by almost 30%, from 13.64 in BASEUNFAV down to 9.78 in UNFAV; the difference being significant (ranksum test, p=.0007, two-sided). Our favorable manipulation does not have the same power. Our data suggest that "fixing windows" in an uncooperative environment is not as straightforward as breaking them in an otherwise cooperative environment. The favorable information increases the mean contribution in the first period only very slightly, from 10.35 in BASEFAV to 10.48 in FAV. This difference is not significant.

Taken together, these two findings provide a possible explanation for the overall difference in cooperation that we observe between treatments. Contributions are path-dependent, i.e., they depend on contributions in the first period. These first contributions, in turn, are apparently manipulable by our very simple mechanism (though the strength of this manipulation seems to be

13

asymmetric). Thus, if people are provided with unfavorable information, i.e., if they observe "broken windows", they start with low cooperation rates and also stay at low cooperation rates in the long run; favorable information at least makes it easier to sustain cooperation in the long run. This suggests that subjects seem to form an initial expectation about how others are likely to behave and base their decision on this expectation. The initial expectations then appear as some kind of self-fulfilling prophecy: The mere pre-game expectation of an uncooperative environment leads to an irrevocable loss of cooperation throughout the game (and vice versa).

PRE-GAME EXPECTATIONS:

To shed more light on the role of expectations, we subsequently ran additional experiments in 2010 at the Econ Lab of the University of Jena, Germany. In these experiments, subjects received exactly the same instructions and had to answer the same control questions as in the games above. However, instead of actually playing the game they were then told that these experiments had been conducted before. Their task was to guess how much those previous participants contributed on average to the common project in the first period (rounded to the next integer). Moreover, subjects had to state how confident they were in this expectation (on a Likert scale ranging from 1 to 7, where 1 means "being very unsure" and 7 means "being very sure" about the belief).[8]

Altogether, 96 subjects participated in these experiments. In the first treatment, subjects received the instructions from the baselines. In the second treatment, they received the instructions of treatment Fav. Finally in the third treatment they received the instructions of treatment Unfav.
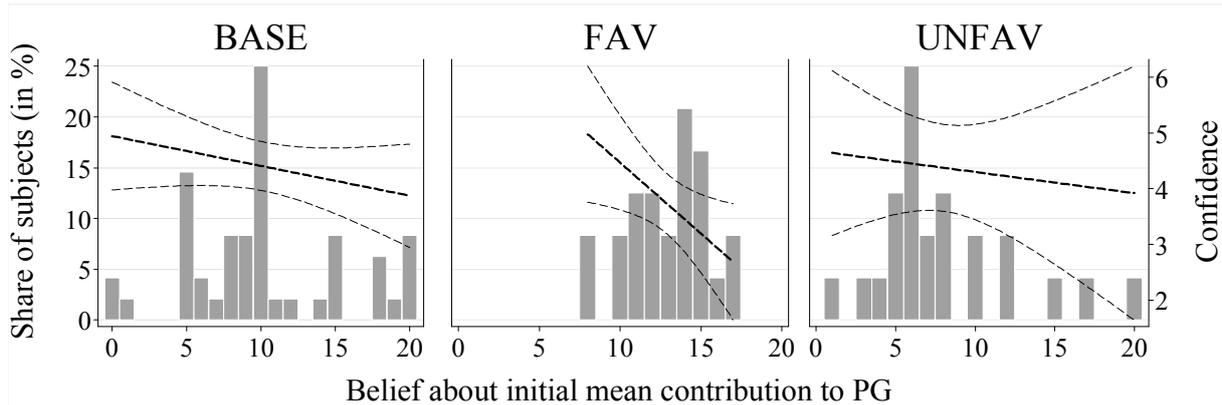
Fig. 6 –Beliefs and Confidence

Notes: BASE has 48 subjects, FAV and UNFAV have 24 subjects respectively. The bars denote the share of subjects in a given treatment expressing a certain belief. The dashed lines represent linear prediction plots for the subjects' self-reported Confidence (with 95% confidence intervals). Confidence is measured on a 7-point Likert scale ranging from 1 ("I am very unsure") to 7 ("I am very sure").

---

8   In the belief elicitation stage, subjects had to guess a whole number between 0 and 20. If the guess was exactly correct, they received 4 Euros. If the guess was wrong, they received 0 Euros. The confidence question was not incentivized. See Appendix 5 for details.

As Fig. 6 shows, the information manipulation has a strong effect on beliefs. In BASE, subjects expect the first period's contribution to be 10.27 on average. The belief drops to 8.04 in UNFAV, i.e. if they know that participants received the unfavorable information sheet. The difference is significant (ranksum test, p=0.0472, two-sided). By contrast, the average belief increases from 10.27 to 12.96 in FAV. Again, the difference is significant (ranksum test, p=0.0054, two-sided).

Mean confidence does not vary significantly across treatments. Interestingly however, as can be seen in Fig. 6, in each treatment confidence is negatively correlated with beliefs. Subjects with a relatively pessimistic view of other's cooperativeness are considerably more confident than those with optimistic beliefs. This pattern is significant and very robust[9] and might give us an indication for why actual contributions react strongly to negative selective information, but only cautiously to positive selective information. The less confident optimistic beliefs might be more prone to external manipulation than the more confident pessimistic beliefs.

Taken together, the additional experiments provide further evidence for the above interpretation of our results from the repeated public-good game. If people face a cooperation problem and observe "broken windows", they strongly expect that others are not likely to cooperate, so they themselves also do not cooperate. By contrast, if "windows are fixed", they believe that others might cooperate, but at the same time acknowledge the cooperation dilemma. Thus, they are not very confident in their expectation. This translates into only intermediate rates of cooperation – though this seemingly suffices to at least sustain cooperation over time.

5   Discussion

In this paper, we have explored whether cooperative behavior reacts to selective pre-play information about other, unrelated groups. To this end, we have used an experimental framework that captures cooperation situations which are sufficiently rich, so that several potentially conflicting behavioral norms might emerge. We find that the aggregate level of cooperativeness is dramatically reduced by giving subjects examples of uncooperative behavior. The opposite intervention was less powerful, but still effective. If participants in an otherwise uncooperative environment saw examples of successful groups, cooperation rates were stabilized, while they quickly decayed otherwise. A likely reason for these findings is that cooperation is path-dependent, i.e., long-run behavior is strongly correlated with initial contribution rates – which, in turn, are a reaction to the pre-play information given to the subjects. The latter might be due to a change in subjects' expectations about others' behavior, as was suggested by our additional experiments.

Our findings have a number of important implications. First, they clearly point to the relevance of pre-game communication – a factor which has only lately started to receive significant attention in the literature. While the existing literature usually focuses on self-chosen cheap talk mes-

---

9   Log-Level OLS regression of ln(confidence) on belief. A 1-point increase in beliefs corresponds on average to a 2 % decrease in confidence (p=0.004). This result is robust to adding treatment dummies, changing the functional form to Level-Level or Log-Log, or running a Tobit.

sages (for an overview, see Crawford 1998), we demonstrate that also exogenously selected, one-way information about other players can alter how players act in subsequent games. In particular the findings of our second series of experiments might be of interest to this literature. They suggest a possible channel through which the observed effects of cheap talk are mediated, namely through the alteration of subjects' pre-game expectations.

Second, and closely related, is our finding that there seems to be a strong asymmetry in the strength of the reaction to the mechanism. By giving subjects examples of uncooperative behavior, the aggregate level of cooperativeness is immediately reduced. By contrast, optimistic foreign experiences need time to unfold their beneficial effect. Bad impressions carry more weight than good ones, so that creating bad behavior is seemingly easier than producing good one – a finding which is also mirrored in previous work on the prevalence of "bad" vs. "good" (cf. Baumeister, Bratslavsky et al. 2001). Thus, one might expect that also other mechanisms which build on the manipulation of beliefs perform differently, depending on the content of the experience as well as on the specific environment in which they occur.

Taken together, our results underline the power and importance of information and experience in shaping cooperative behavior. The bottom line is that observation matters. Interestingly, people do not only learn from what they experience themselves. They also seem to learn "vicariously", by observing others, or by seeing the results (Bandura 1977). The effect is even present when participants are told that the information they are receiving is selective. This connects our study with the growing literature on social learning (for a recent meta-study see Weizsäcker 2010). Our subjects' behavior critically depends on pre-game expectations, which we show to be easy to deteriorate in a complex setting – simply by providing subjects the opportunity of vicarious learning.

This suggests that, in appropriate circumstances, impression management might indeed be a feasible tool to avert, or at least to mitigate, the danger of social dilemmas. Of course this is a paternalistic intervention. But note that conditional cooperators need not even be deceived for the intervention to be successful. All that is required is that they expect the manipulation to matter for a sufficient fraction of the remaining members of their group, be that because they are deceived, or because they are skeptical themselves, but willing to give cooperation a try since the intervention gives all of them one and the same informational starting point.

For policy makers these findings represent both a chance and a peril. If they do not manage to repair broken windows quickly, both literally and metaphorically speaking, chances are a vicious cycle starts. By contrast, if they can induce some to lead others by their socially beneficial example, this strategy may well work. In particular, policy makers would want to prevent (perhaps wrong) pessimistic beliefs from spreading. In any case, we show that home-grown expectations must not be disregarded in order to attain socially desirable outcomes. Consequently in situations where the success of a law depends on the willingness of individual citizens to cooperate – for instance in areas like waste separation and sustainable water use – government might want to

consider a PR campaign in order to create a general atmosphere of cooperativeness within the population.

Managing first impressions might certainly be less effective if the large majority of addressees know better. The intervention requires a sufficient degree of uncertainty. Yet in political reality, quite a few public goods are characterized by deep conceptual and factual uncertainty. Problems like climate change are heavily contested among scientists and not well understood by many. Addressees have to trust expertise. If in the eyes of addressees the underlying social problem is opaque, they are also likely to be uncertain how others will react to it. Most importantly, addressees face behavioral uncertainty whenever they newly enter a community. They do not know local mores, nor do they know how determined the group is when it comes to enforcing them.

The results in this paper should, of course, not be taken as arguments against the importance of elaborated, incentives-altering mechanisms in general. For instance, the effectiveness of our very simple mechanism seems to be asymmetric, since it has a hard time to increase cooperation in an otherwise uncooperative environment. Still, our findings suggest that minimal interventions *can* have a strong behavioral effect. Future research could try to explore the interaction between such simple and other, more complex and intrusive mechanisms. Besides, it might also be informative to study in more detail if certain player "types" (like selfish actors, conditional cooperators, or altruists) are affected differently by mechanisms in the spirit of the one presented here. It is even conceivable that such interventions change the preferences of certain players. Another important, though challenging, next step would certainly be to move into the field and to investigate the effectiveness of means to foster cooperation in a more natural setup. Extending on our work allows for studying these and other interesting aspects in the future.

References

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

ANDREONI, JAMES and RAGAN PETRIE (2004). "Public Goods Experiments without Confidentiality. A Glimpse into Fund-raising." <u>Journal of Public Economics</u> **88**: 1605-1623.

BANDURA, ALBERT (1977). <u>Social Learning Theory</u>. Englewood Cliffs, Prenctice Hall.

BAUMEISTER, ROY, ELLEN BRATSLAVSKY, CATRIN FINKENAUER and KATHLEEN D. VOHS (2001). "Bad is Stronger than Good." <u>Review of General Psychology</u> **5**: 323-370.

BECKENKAMP, MARTIN, CHRISTOPH ENGEL, ANDREAS GLÖCKNER, BERND IRLENBUSCH, HEIKE HENNIG-SCHMIDT, SEBASTIAN KUBE, MICHAEL KURSCHILGEN, ALEXANDER MORELL, ANDREAS NICKLISCH, HANS-THEO NORMANN and EMANUEL TOWFIGH (2009). Beware of Broken Windows! First Impressions in Public-good Experiments, mimeo.

BLUME, ANDREAS and ANDREAS ORTMANN (2007). "The Effects of Costless Pre-Play Communication. Experimental Evidence from Games with Pareto-Ranked Equilibria." <u>Journal of Economic Theory</u> **132**: 274-290.

BRAGA, ANTHONY A., DAVID L. WEISBURD, ELIN J. WARING, LORRAINE GREEN MAZEROLLE, WILLIAM SPELMAN and FRANCIS GAJEWSKI (1999). "Problem-Oriented Policing in Violent Crime Places. A Randomized Controlled Experiment." <u>Criminology</u> **37**: 541-580.

BREKKE, KJELL ARNE, KAREN EVELYN HAUGE, JO THORI LIND and KARINE NYBORG (2009). Playing with the Good Guys: A Public Good Game with Endogenous Group Formation, mimeo.

BURLANDO, ROBERTO M. and FRANCESCO GUALA (2005). "Heterogeneous Agents in Publilc Goods Experiments." <u>Experimental Economics</u> **8**: 35-54.

CAMERON, ADRIAN COLIN and P. K. TRIVEDI (2005). <u>Microeconometrics. Methods and Applications</u>. Cambridge ; New York, Cambridge University Press.

CIALDINI, ROBERT B., RAYMOND R. RENO and CARL A. KALLGREN (1990). "A Focus Theory of Normative Conduct. Recycling the Concept of Norms to Reduce Littering in Public Places." <u>Journal of Personality and Social Psychology</u> **58**: 1015-1026.

CINYABUGUMA, MATTHIAS, TALBOT PAGE and LOUIS PUTTERMAN (2005). "Cooperation under the Threat of Expulsion in a Public Goods Experiment." <u>Journal of Public Economics</u> **89**: 1421-1435.

CORAZZINI, LUCCA, SEBASTIAN KUBE, MICHEL ANDRÉ MARÉCHAL and ANTONIO NICOLO (2010). Elections and Deceptions. Theory and Experimental Evidence

CRAWFORD, VINCENT (1998). "A Survey of Experiments on Communication via Cheap Talk." Journal of Economic Theory **78**(2): 286-298.

CROSON, RACHEL T.A., ENRIQUE FATAS and TIBOR NEUGEBAUER (2008). The Effect of Excludability on Team Production, mimeo.

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

FALK, ARMIN and URS FISCHBACHER (2006). "A Theory of Reciprocity." Games and Economic Behavior **54**: 293-315.

FALKINGER, JOSEF, ERNST FEHR, SIMON GÄCHTER and RUDOLF WINTER-EBMER (2000). "A Simple Mechanism for the Efficient Provision of Public Goods. Experimental Evidence." American Economic Review **90**: 247-264.

FEHR, ERNST and SIMON GÄCHTER (2000). "Cooperation and Punishment in Public Goods Experiments." American Economic Review **90**: 980-994.

FEHR, ERNST and SIMON GÄCHTER (2002). "Altruistic Punishment in Humans." Nature **415**: 137-140.

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 and ERNST FEHR (2001). "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." Economics Letters **71**: 397-404.

FORSYTHE, ROBERT, ROGER B. MYERSON, THOMAS RIETZ and ROBERT J. WEBER (1993). "An Experiment on Coordination in Multi-Candidate Elections. The Importance of Polls and Election Histories." Social Choice & Welfare **10**: 223-247.

FORSYTHE, ROBERT, THOMAS RIETZ, ROGER B. MYERSON and ROBERT J. WEBER (1996). "An Experimental Study of Voting Rules and Polls in Three-Candidate Elections." International Journal of Game Theory **25**: 355-383.

GÄCHTER, SIMON and CHRISTIAN THÖNI (2007). "Social Learning and Voluntary Cooperation Among Like-Minded People." Journal of the European Economic Association **3**: 303-314.

GALBIATI, ROBERTO and PIETRO VERTOVA (2008). "Law and Behaviours in Social Dilemmas. Testing the Effect of Obligations on Commitment." Games and Economic Behavior **64**: 146-170.

GLÖCKNER, ANDREAS, BERND IRLENBUSCH, ANDREAS NICKLISCH and HANS-THEO NORMANN (2010). "Leading With(out) Sacrifice? A Public-Goods Experiment with a Privileged Player." Economic Inquiry ***: ***.

GOERG, SEBASTIAN J. and GARY WALKOWITZ (2010). "On the Prevalence of Framing Effects Across Subject-Pools in a Two-Person Cooperation Game." Journal of Economic Psychology ***: ***.

GREINER, BEN (2004). An Online Recruiting System for Economic Experiments. Forschung und wissenschaftliches Rechnen 2003. Kurt Kremer und Volker Macho. Göttingen: 79-93.

GUNNTHORSDOTIR, ANNA, DANIEL HOUSER and KEVIN MCCABE (2007). "Disposition, History and Contributions in Public Goods Experiments." Journal of Economic Behavior & Organization 62: 304-315.

GÜRERK, ÖZGÜR, BERND IRLENBUSCH and BETTINA ROCKENBACH (2006). "The Competitive Advantage of Sanctioning Institutions." Science 312: 108-111.

HERRMANN, BENEDIKT, CHRISTIAN THÖNI and SIMON GÄCHTER (2008). "Antisocial Punishment Across Societies." Science 319: 1362-1367.

ISAAC, R. MARK, JAMES M. WALKER and ARLINGTON W. WILLIAMS (1994). "Group Size and the Voluntary Provision of Public Goods. Experimental Evidence Utilizing Large Groups." Journal of Public Economics 54: 1-36.

KEIZER, KEES, SIEGWART LINDENBERG and LINDA STEG (2008). The Spreading of Disorder. Science. 322: 1681-1685.

KLOR, ESTEBAN and EYAL WINTER (2007). "The Welfare Effects of Public Opinion Polls." International Journal of Game Theory 35: 379-394.

KOGAN, SHIMON, ANTHONY M. KWASNICA and ROBERTO A. WEBER (2010). "Coordination in the Presence of Asset Markets." American Economic Review ***: ***.

KUBE, SEBASTIAN and CHRISTIAN TRAXLER (2010). "The Interaction of Legal and Social Norm Enforcement." Journal of Public Economic Theory ***: ***.

MAMMEN, ENNO (1993). "Bootstrap and Wild Bootstrap for High Dimensional Linear Models." Annals of Statistics 21: 255-285.

MENGEL, FRIEDRIKE (2007). "The Evolution of Function-Value Traits for Conditional Cooperation." Journal of Theoretical Biology 245: 564-575.

NIKIFORAKIS, NIKOS S. (2008). "Punishment and Counter-Punishment in Public Good Games: Can We Really Govern Ourselves?" Journal of Public Economics 92: 91-112.

OSTROM, ELINOR, JAMES M. WALKER and ROY GARDNER (1992). "Covenants with and without Sword. Self-Governance is Possible." American Political Science Review **40**: 309-317.

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

RAMOS, JOAO and BENNO TORGLER (2010). Are Academics Messy? Testing the Broken Windows Theory with a Field Experiment in the Work Environment http://www.bepress.com/cgi/viewcontent.cgi?article=1505&context=feem.

WEBER, ROBERTO A. (2006). "Managing Growth to Achieve Efficient Coordination in Large Groups." American Economic Review **96**: 114-126.

WEIZSÄCKER, GEORG (2010). "Do We Follow Others When We Should? A Simple Test of Rational Expectations." American Economic Review *******.

WILSON, JAMES Q. and GEORGE L. KELLING (1982). "Police and Neighborhood Safety. Broken Windows." Atlantic Monthly **127**: 29-38.

WU, C.F.J. (1986). "Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis." Annals of Statistics **14**: 1261-1295.

ZIMBARDO, PHILIP (1969). "The Human Choice. Individuation, Reason, and Order versus Deindividuation, Impulse, and Chaos." Nebraska Symposium on Motivation **17**: 237-307.

ZIMBARDO, PHILIP G. and EBBE B. EBBESEN (1969). Influencing Attitudes and Changing Behavior. A Basic Introduction to Relevant Methodology, Theory, and Applications. Reading, Mass.,, Addison-Wesley Pub. Co.

Appendix 1: Instructions Baseline

You are now taking part in an economic experiment. If you read the following instructions carefully, you can, depending on your decisions, earn a considerable amount of money. It is therefore important that you take your time to understand the instructions.

The instructions which we have distributed to you are for your private information. **Please do not communicate with the other participants during the experiment.** Should you have any questions please ask us.

During the experiment we shall not speak of Pounds, but of Experimental Currency Units (ECU). Your entire earnings will be calculated in ECUs. At the end of the experiment the total amount of ECUs you have earned will be converted to Pounds at the rate of **1 ECU = 4 p** and will be immediately paid to you in cash.

At the beginning of the experiment the participants will be randomly divided into groups of four. You will therefore be in a group with 3 other participants. **The composition of each group will remain the same throughout the experiment.** The experiment lasts 10 periods and each period is divided into 3 stages.

The 1st stage:

At the beginning of each of the 10 periods each participant will receive 20 ECUs. In the following, we shall refer to this amount as the "endowment". In the 1st stage, your task is to decide how to use your endowment. You have to decide how many of the 20 ECUs you want to contribute to a project (from 0 to 20) and how many of them to keep for yourself. The consequences of your decision are explained in detail below.

Once all the players have decided their contribution to the project you will be informed about, the group's total contribution, your income from the project and your payoff in this period. Your payoff in each period is calculated using the following simple formula. Again, if you have any difficulties do not hesitate to ask us.

Income from the 1st stage	**=**	Endowment of ECUs	**−**	Your contribution to the Project	**+**	0.4*Total contribution to the Project

This formula shows that your 1st stage income consists of two parts:
1) The ECUs which you have kept for yourself (endowment – contribution)
2) The income from the project, which equals to the 40% of the group's total contribution.

The income of each group member from the project is calculated in the same way. This means that each group member receives the same income from the project. Suppose the sum of the contributions of all group members are 60 ECUs. In this case, each member of the group receives an income from the project of: 0.4*60=24 ECUs. If the total contribution to the project is 9 points, then each member of the group receives an income of: 0.4*9=3.6 ECUs from the project.

You always have the option of keeping the ECUs for yourself or contributing them to the project. Each ECU that you keep raises your end of period income by 1 ECU. Supposing you contributed this point to the project instead, then the total contribution to the project would rise by 1 ECUs. Your income from the project would thus rise by 0.4*1=0.4 ECUs. However, the income of the other group members would also rise by 0.4 ECUs each, so that the total income of the group from the project would be 1.6 points. Your contribution to the project therefore also raises the income of the other group members. On the other hand you also earn an income for each point contributed by the other members to the project. In particular, for each point contributed by any member you earn 0.4 ECUs.

In addition to the 20 ECUs per period, each participant receives a one-off lump sum payment of 25 ECUs at the beginning of this part. This one-off payment can be used to pay for eventual losses during the experiment. **However, you can always evade losses <u>with certainty</u> through your own decisions.** Note that this lump sum payment will not be used to calculate the income from the period. It will only be added to your total income from all the periods at the very end.

The 2nd stage:

At the 2nd stage you will be informed how much each group member contributed individually to the project at the 1st stage. At this stage you can **reduce or leave equal** the income of **each** member of your group by **distributing points**. The other group members can also reduce your income if they wish to.

If you choose 0 points for a particular group member, you do not change his or her income. However if you give a member 1 point you reduce his or her income by 10 percent. If you give a member 2 points you reduce his or her income by 20 percent, etc. The amount of points you distribute to each member determines, therefore, how much you reduce their income from the 1st stage. If one player receives in total 4 points his income will be reduced by 40% and if he receives **10 or more** his income from the 1st stage will be reduced by 100%.

If you distribute points you have costs in ECUs, which depend on the amount of points you distribute. You can distribute between 0 and 10 points to each group member. The more points you give **to any** group member, the higher your costs. Your total costs are equal to the **sum of the costs of distributing points to each of the other three group members**. The following table illustrates the relation between distributed points to **each** group member and the cost of doing so in ECUs.

Points	0	1	2	3	4	5	6	7	8	9	10
Cost of points per person	0	1	2	4	6	9	12	16	20	25	30

Example: Supposing you give 2 points to player 1 this costs you 2 ECUs; if you also give 8 points to player 3 this costs you a further 20 ECUs; and if you give 0 points to the last group member this has no additional cost for you. In this case, your total costs of distributing points would be 22 ECUs (2+20+0) and not 30 ECUs.

The following equation summarizes the previous information. Your total income from the two stages is calculated as follows:

Income at the end of the 2nd stage	=	Income from the 1st stage	*	[(10 – received points)/10]	–	Costs of distributed points

Please note that your income in ECUs at the end of the 2nd and the 3rd stage can be negative, if the costs of your points distributed exceeds your (possibly reduced) income from the 1st stage. **You can however evade such losses with certainty through your own decisions.** Should your income become zero or negative at the end of the 2nd stage you will not be able to continue to the 3rd stage. If your income becomes zero or negative at the end of the 3rd stage you can simply use your 25 ECUs that we gave you in the beginning in order to pay this off.

The 3rd stage:
In the 3rd and final stage, after being informed of the points that the other group members assigned to you, you will be given one last opportunity of assigning points back to the other participants, thus reducing their income. We shall call these points "counter-points". **You will only be able to assign counter-points to participants who assigned points to you during the 2nd stage**

The costs of assigning points, as well as the income reduction caused by each point remain the same as before. The following table shows you how to calculate the costs for assigning counter-points.

Costs for assigning **counter-points** to one specific group member in the 3rd stage											
Points that you already assigned to one specific group member in the 2nd stage	**Counter-points** you assign to that same group member in the 3rd stage										
	0	1	2	3	4	5	6	7	8	9	10
0	0	1	2	4	6	9	12	16	20	25	30
1	0	1	3	5	8	11	15	19	24	29	
2	0	2	4	7	10	14	18	23	28		
3	0	2	5	8	12	16	21	26			
4	0	3	6	10	14	19	24				
5	0	3	7	11	16	21					
6	0	4	8	13	18						
7	0	4	9	14							
8	0	5	10								
9	0	5									
10	0										

Example: if you distribute 2 points in the 2nd stage to player 1 you have a cost of 2 ECUs If in the 3rd stage you decide to distribute 3 counter-points to player 1, a further 7 ECUs are added to your cost.

Your income after the 3rd stage (= period income) is therefore calculated as follows:

Income at the end of the 3rd stage	=	Income from the 2nd stage	*	[(10 – received counter-points)/10]	– Cost of distributed counter-points

If you have any further questions please raise your hand and one of the supervisors

Appendix 2: Control Questionnaire

1. Each group member has an endowment of 20 ECUs. Nobody (including you) contributes any ECUs to the project. What is:
   a. Your income at the end of the first stage? ………
   b. The income of the other group members?………

2. Each group member has an endowment of 20 ECUs. You contribute 20 ECUs to the project. All other group members contribute 20 ECUs each to the project. What is:
   a. Your income at the end of the first stage? ………
   b. The income of the other group members?………

3. Each group member has an endowment of 20 ECUs. The other three group members contribute together a total of 30 ECUs to the project. What is:
   a. Your income at the end of the first stage if you contribute 0 ECUs to the project? ……….
   b. Your income at the end of the first stage if you contribute 15 ECUs to the project? ……….

4. Each group member has an endowment of 20 ECUs. You contribute 8 ECUs to the project. What is:
   a. Your income at the end of the first stage if the other group members together contribute a further total of 7 ECUs to the project?...............
   b. Your income at the end of the first stage if the other group members together contribute a further total of 22 ECUs to the project?...............

5. At the second stage you distribute the following points to your three other group members: 9, 5, 0. What are the total costs of your distributed points?….

6. What are your costs if you distribute 0 points? ……

7. By how many percent will your income from the first stage be reduced when you receive from the other group members a total of:
   a. 0 points? …
   b. 4 points? ...
   c. 15 points? …

8. At the second stage you distribute the following points to your three other group members: 2, 2, 0. In the third stage you distribute the following points to your three other group members: 1, 1, 1. What are the total costs of your distributed points?….

9. By how many per cent is your second stage income reduced, if you have received the sum of 3 counter-points from the other group members in stage 3?

Appendix 3: Additional Sheet FAV Treatment

Additional information for today's experiment

This experiment has already been run at two laboratories in Bonn (Germany) and London (UK). The following figure shows you the contribution behaviour in four selected groups from both locations.
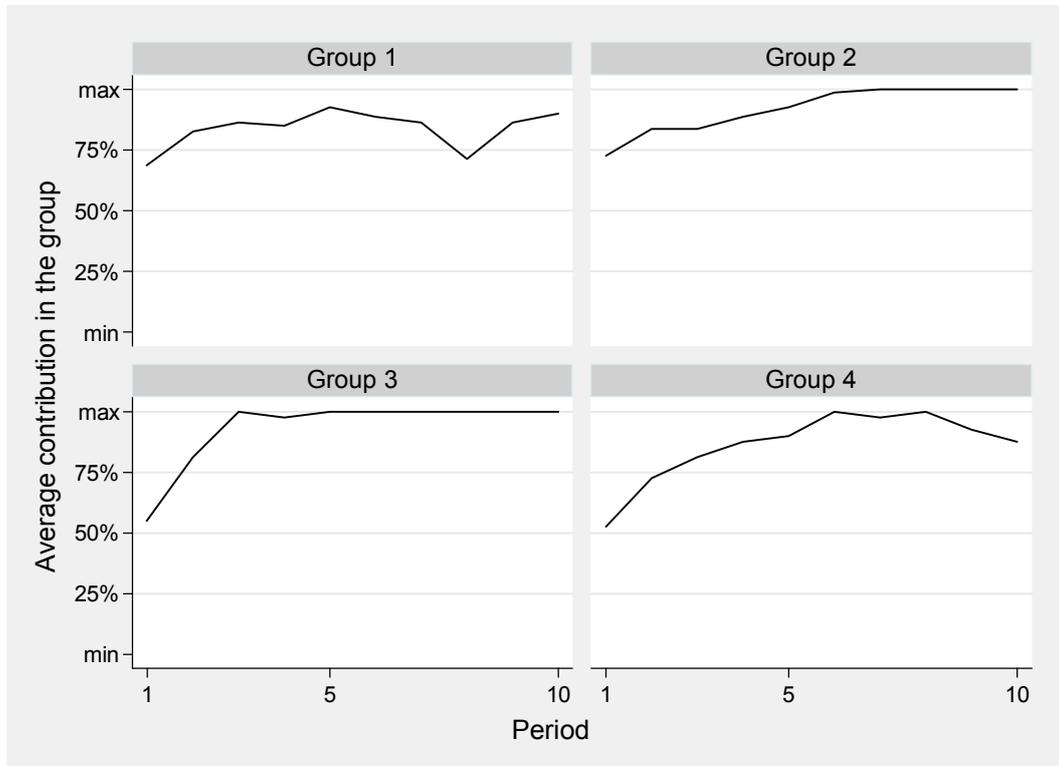
In these selected graphs you see how much the 4 group members contributed on average to the group project. In the selected groups 1 and 2 the contributions are high right from the beginning. In the selected groups 3 and 4, average contribution starts somewhat lower but rises over the course of the experiment.
Some additional numbers from the experiments in Bonn and London:

a) In the first round of the two previous experiments, 58% to 74% of the contributions were between 10 and 20.
b) In the last round, up to 53% of the contributions were equal to 20.
c) Players, whose income was reduced in stage 2, only used counter-points in stage 3 in 29% to 41% of all possible cases.

All these graphs and numbers are meant to give you some orientation concerning the type of situation you might be in during the experiment and what you might possibly expect from the other members of your group.

Appendix 4: Additional Sheet UNFAV Treatment

Additional information for today's experiment

This experiment has already been run at two laboratories in Bonn (Germany) and London (UK). The following figure shows you the contribution behaviour in four selected groups from both locations.

In these selected graphs you see how much the 4 group members contributed on average to the group project. In the selected groups 1 and 2 the contributions are rather low and display a decrease over the course of the experiment. In the selected groups 3 and 4, average contributions are low right from the beginning
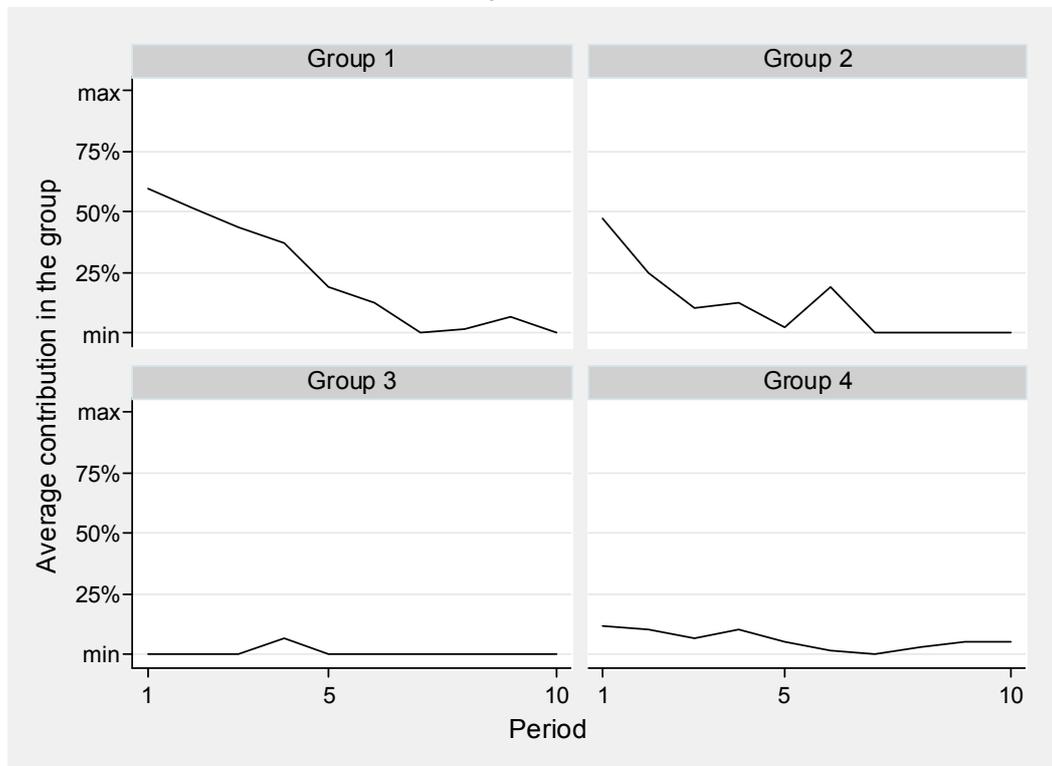Some additional numbers from the experiments in Bonn and London:

a) In the first round of the two previous experiments, 37% to 55% of the contributions were between 0 and 10.
b) In the last round, up to 52% of the contributions were equal to 0.
c) Players, whose income was reduced in stage 2, only used counter-points in stage 3 to reduce the income of those who had given them points in stage 2. Every time they used counter-points they reduced the income of those who had given them points in stage 2 by 18% to 25%.

All these graphs and numbers are meant to give you some orientation concerning the type of situation you might be in during the experiment and what you might possibly expect from the other members of your group.

Appendix 5: Belief Elicitation

As noted in the main text, subjects in the belief elicitation had to read the same instructions and answer the same control questions as the participants in the contribution treatments. However, once ztree started subjects were told the following on their screens:

You have just read the instructions and answered the control questions. This game has already been played in other laboratories. The participants of those experiments read the same instructions as you and answered the same control questions. You are not going to play this game today. Instead, your task is to answer the following question:

What do you think, how much did those previous participants contribute on average in the first round to the project?

Your guess………

We have rounded the average down to the next whole number, i.e. the correct answer is a whole number between 0 and 20. If you guess is exactly correct, you will receive 4 Euros. If your guess is wrong, you will receive 0 Euros.

Appendix 6: Additional Data Analysis

Table A1: Summary Statistics of Punishment and Counter-punishment across Treatments

		BASEFAV	BASEUNFAV	FAV	UNFAV
(1)	Punishment	0.65	0.35	0.92	0.51
(2)	Counter-punishment	0.17	0.11	0.21	0.23
(3)	Counter-punishment (conditional)	0.72	0.78	0.59	1.06

Notes: All figures are means of (counter-)punishment points given per period and subject. (3) is conditional on subjects being able to allocate counter-punishment. Subjects could only allocate counter-punishment in stage 3 if they had been punished in stage 2.

Parametric Analysis

We have participants interacting over 10 announced periods in fixed groups of four. Moreover all our data is right-censored since many participants contribute their entire endowment of 20. Most of our data is also left-censored in that a substantial fraction has kept the entire endowment. The ideal functional form would therefore be a mixed Tobit model. Unfortunately, such a model is not available. Random effects Tobit would ignore that observations within groups are not independent. A linear mixed model would ignore that observations are censored. We therefore estimate Tobit models, with lower level 0 and upper level 20, and correct standard errors via bootstrapping, with resampling at the level of entire groups. All models work with 50 repetitions, which is known to be sufficient for correcting standard errors.

In the comparison between BASEUNFAV and UNFAV, a non-parametric bootstrap (resampling pairs of dependent and independent variables) yields significant effects. We therefore do not report results from a (semi-)parametric bootstrap resampling residuals (for background see Cameron and Trivedi 2005:360 f.), nor from „wild" bootstrapping, which randomly assigns error terms using the procedure developed by (Wu 1986; Mammen 1993), using the parameters proposed by (Cameron and Trivedi 2005:377); unsurprisingly, showing significance becomes even easier if one relies on these methods. By contrast, these more advanced methods make a difference in the comparison between BASEFAV and FAV, which is why we report all three estimates. Since we only use bootstrapping for correcting standard errors for the dependence of observations, parametric bootstrapping is justified.

In the comparison of BASEUNFAV with UNFAV, in all models we find a strong and highly significant negative effect of the belief manipulation. There is no time trend, and no interaction between time and treatment. By contrast, punishment has a highly significant negative main effect: those who have been punished in the previous period still contribute less than the group average in the subsequent period. This holds irrespective of the belief manipulation, as shown by the insignificance of the interaction effect.

Table A 2: Comparing BASEUNFAV with UNFAV						
	model 1	model 2	model 3	model 4	model 5	model 6
unfav	-10.708 (.001)	-10.709 (.001)	-9.275 (.001)	-10.972 (.001)	-11.583 (.004)	-13.768 (.005)
period		.027 (.896)	.163 (.568)	-.421 (.116)	-.474 (.215)	-.615 (.184)
unfav*period			-.263 (.453)		.102 (.796)	.304 (.498)
lagged received punishment (lrecpun)				-1.947 (.013)	-1.949 (.013)	-3.400 (.006)
unfav*lrecpun						2.273 (.172)
cons	22.104 (<.001)	21.957 (<.001)	21.218 (<.001)	26.022 (<.001)	26.347 (<.001)	27.723 (<.001)
depvar: contribution. Tobit, with lower level 0 and upper level 20. Models correct for the dependence of observations per individual and within groups by bootstrapping, with random draws of entire groups. A non-parametric bootstrap is administered, i.e. pairs of dependent and their independent variables are randomly selected. p-values in parenthesis. N = 1152. 95 left censored, 494 right censored observations.						

In the comparison of BASEFAV with FAV we get a different picture. We have a positive main effect of the belief manipulation in all models. But it is significant only with wild bootstrapping (models 1, 2, 4), and weakly significant with model 4 (which controls for lagged received punishment). There is, however, a significant interaction between the treatment and the time trend in model 3. It also is present in wild bootstrapping in models 5 and 6 at conventional significance levels, and it is weakly significant with alternative bootstrapping methodologies. This indicates that the belief manipulation stops contributions from quickly eroding. In models 4 and 5, we also find the negative effect of lagged received punishment, as in the comparison of BASEUNFAV with UNFAV.

Table A 2: Comparing BASEFAV with FAV						
	model 1	model 2	model 3	model 4	model 5	model 6
fav	6.151	6.126	1.071	7.275	3.114	2.612
	(.122)	(.123)	(.789)	(.085)	(.502)	(.662)
	(.211)	(.181)	(.745)	(.074)	(.332)	(.433)
	(.004)	(.006)	(.683)	(.001)	(.286)	(.446)
period		-.584	-1.114	-.939	-1.335	-1.348
		(.010)	(.001)	(<.001)	(<.001)	(.001)
		(.001)	(<.001)	(<.001)	(<.001)	(<.001)
		(<.001)	(<.001)	(<.001)	(<.001)	(<.001)
fav*period			.925		.695	.726
			(.015)		(.092)	(.101)
			(.019)		(.081)	(.072)
			(<.001)		(.014)	(.018)
lagged received punish-ment				-1.395	-1.369	-1.610
				(.038)	(.046)	(.222)
				(.024)	(.021)	(.101)
				(<.001)	(<.001)	(.003)
fav*lrecpun						.405
						(.791)
						(.709)
						(.560)
cons	7.714	10.906	13.790	13.867	16.205	16.444
	(.023)	(.002)	(<.001)	(.002)	(.001)	(.002)
	(.002)	(<.001)	(<.001)	(<.001)	(<.001)	(<.001)
	(<.001)	(<.001)	(<.001)	(<.001)	(<.001)	(<.001)

depvar: contribution. Tobit, with lower level 0 and upper level 20. Models correct for the dependence of observations per individual and within groups by bootstrapping, with random draws of entire groups. We compare three different bootstrapping techniques, with p-values in parenthesis and in the following order: non-parametric; parametric over residuals; wild bootstrap. N = 1008. 214 left censored, 277 right censored observations.

Preprints 2011

2011/04: Petersen N., The Role of Consent and Uncertainty in the Formation of Customary International Law

        forthcoming in: Reexamining Customary International Law, Cambridge, Cambridge University Press, In Press.

2011/03: Petersen N., Antitrust Law and the Promotion of Democracy and Economic Growth

2011/02: Engel C., Besonderes Verwaltungsrecht und ökonomische Theorie

2011/01: Engel C., Zhurakhovska L., Oligopoly as a Socially Embedded Dilemma. An Experiment

Preprints 2010

2010/48: Engel C., An Experimental Contribution to the Revision of the Guidelines on Research and Development Agreements

2010/47: Bade S., Pareto-Optimal Matching Allocation Mechanisms for Boundedly Rational Agents

2010/46: Bade S., Matching Allocation Problems with Endogenous Information Acquisition

2010/45: Bierbrauer F., Boyer P. C., Political competition and Mirrleesian income taxation: A first pass

2010/44: Enders Z., Hakenes H., On the Existence and Prevention of Asset Price Bubbles

2010/43: Weinschenk P., Increasing Workload in a Stochastic Environment

2010/42: Admati A. R., DeMarzo P. M., Hellwig M., Pfleiderer P., Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is Not Expensive

2010/41: Körner T., Schnabel I., Public Ownership of Banks and Economic Growth – The Role of Heterogeneity

2010/40: Ganuza J. J., Jansen J., Too Much Information Sharing? Welfare Effects of Sharing Acquired Cost Information in Oligopoly

2010/39: Weinschenk P., Moral Hazard and Ambiguity

2010/38: Herweg F., Müller D., Weinschenk P., Binary Payment Schemes: Moral Hazard and Loss Aversion

        published in: American Economic Review, vol. 100, pp. 2451-2477, 2010.

2010/37: Glöckner A., Engel C., Role Induced Bias in Court: An Experimental Analysis

2010/36: Dang T. V., Hakenes H., Information Disclosure, Intertemporal Risk Sharing, and Asset Prices

2010/35: Traxler C., Burhop C., Poverty and crime in 19th century Germany: A reassessment

2010/34: Dickert S., Herbig B., Glöckner A., Gansen C., Portack R., The More the Better? Effects of Training and Information Amount in Legal Judgments

2010/33: Towfigh E. V., Glöckner A., Game Over: Empirical Support for Soccer Bets Regulation

2010/32: Glöckner A., Towfigh E. V., Geschicktes Glücksspiel. Die Sportwette als Grenzfall des Glücksspielrechts

        published in: JuristenZeitung, vol. 21, pp. 1027-1035, 2010.

2010/31: Hellwig M., Capital Regulation after the Crisis: Business as Usual?

2010/30: Prantl S., The Impact of Firm Entry Regulation on Long-living Entrants

        forthcoming in: Small Business Economics, In Press.

2010/29: Engel C., Kurschilgen M., Fairness Ex Ante & Ex Post – The Benefits of Renegotiation in Media Markets

2010/28: Goerg S. J., Walkowitz G., On the Prevalence of Framing Effects Across Subject-Pools in a Two- Person Cooperation Game

        published in: Journal of Economic Psychology, vol. 31, pp. 849-859, 2010.

2010/27: Chmura T., Engel C., Englerth M., Pitz T., At the Mercy of the Prisoner Next Door. Using an Experimental Measure of Selfishness as a Criminological Tool

2010/26: Weinschenk P., Skill Formation under Incomplete Information

2010/25: Brilon S., Job Assignment with Multivariate Skills

2010/24: Lang M., Wambach A., The fog of fraud – mitigating fraud by strategic ambiguity

2010/23: Albanese G., Sorge M. M., The Role of the Judiciary in the Public Decision Making Process

2010/22: Bade S., Electoral Competition with Uncertainty Averse Parties

2010/21: Leifeld P., Haunss S., A Comparison between Political Claims Analysis and Discourse Network Analysis: The Case of Software Patents in the European Union

2010/20: Burhop C., Gelman S., Transaction costs, liquidity and expected returns at the Berlin Stock Exchange, 1892-1913

2010/19: Hellwig M., Finanzkrise und Reformbedarf [The Financial Crisis and Regulatory Reform]

2010/18: Bierbrauer F., Optimal Income Taxation and Public-Goods Provision with Preference and Productivity Shocks

2010/17: Weinschenk P., Entry and Incumbent Innovation

2010/16: Bierbrauer F., Boyer P. C., The Pareto-Frontier in a simple Mirrleesian model of income taxation

2010/15: Grechenig K., Sekyra M., No Derivative Shareholder Suits in Europe – A Model of Percentage Limits and Collusion

2010/14: Bierbrauer F., On the optimality of optimal income taxation

2010/13: Engel C., An Experimental Contribution to the Theory of Customary (International) Law

2010/12: Leifeld P., Schneider V., Institutional communication revisited: Preferences, opportunity structures and scientic expertise in policy networks

2010/11: Grechenig K., Nicklisch A., Thöni C., Punishment despite Reasonable Doubt – A Public Goods Experiment with Uncertainty over Contributions

2010/10: Petersen N., Braucht die Rechtswissenschaft eine empirische Wende?

2010/09: Bade S., Ambiguous Act Equilibria

2010/08: Lehmann S., Volckart O., The Political Economy of Agricultural Protection: Sweden 1887

2010/07: Engel C., Dictator Games: A Meta Study

2010/06: Engel C., Irlenbusch B., Turning the Lab into Jeremy Bentham's Panopticon. The Effect of Punishment on Offenders and Non-Offenders

2010/05: Gropp R., Hakenes H., Schnabel I., Competition, Risk-Shifting, and Public Bail-out Policies

2010/04: Slemrod J., Traxler C., Optimal observability in a linear income tax

2010/03: Baumann, Florian; Friehe, Tim; Grechenig, Kristoffel: Switching Consumers and Product Liability: On the Optimality of Incomplete Strict Liability

2010/02: Bierbrauer F., Hellwig M., Public-Good Provision in a Large Economy

2010/01: Bierbrauer F., Incomplete contracts and excludable public goods