

Social Dilemmas and Shame-based Sanctions: Experimental results from rural Zimbabwe

WPS/2001.11

Abigail Barr

**Centre for the Study of African Economies
University of Oxford**

June 2001

Acknowledgements: Many thanks to my research assistants Michael Shambare, Nyaradzo Dzobo-Shayanewako, Chipso Mukova, and Brighton Denga for their perseverance. Thank you also to Bill Kinsey for introducing me to the villages where the experiments were conducted and to the members of the Cross-Cultural Experimental Economics group, the CSAE, and the Department of Economics at Oxford University, who offered invaluable comments. Finally, I owe a huge debt of thanks to the villagers who took part in the study for their patience and good humour. All remaining errors are my own. This research was funded by the Department for International Development through ESCOR grant R7650.

Abstract: Using two economic experiments I investigate how a sample of rural communities in Zimbabwe approach social dilemmas. When provided with an opportunity to impose sanctions in the context of a public goods game, fourteen out of eighteen communities achieved higher levels of cooperation. In thirteen communities the imposition of shame-based sanctions in the form of light-hearted criticism was observed. The resulting data revealed that: both non-cooperators and cooperators were criticised; community members cared about what their neighbours thought of them and made adjustments to their behaviour accordingly; the overall pattern rather than individual experiences of criticism affected subsequent behaviour; those who made low contributions and witnessed the criticism of others who made similar contributions, made higher contributions subsequently; while those who experienced such criticism first-hand made significantly smaller adjustments to their behaviour; those who made high contributions and witnessed the criticism of others who made similar contributions, made lower contributions subsequently; and to the extent that an opportunity to criticise passed by unexploited subsequent levels of cooperation were reduced.

1. Introduction

In the OECD countries of the North and West, markets are well developed and governments relatively efficient in applying interventions designed to address their failures. As a result, the set of social dilemmas that communities must solve is relatively limited. In developing economies the opposite is true: missing markets and poorly performing governments leave communities with a much wider range of social dilemmas. Common pool resources (CPR) need to be protected from over use, infrastructure must be maintained, informal risk-sharing and credit arrangements are required if incomes are to be smoothed, and economies of scale in service delivery, production and marketing are there to be exploited. A community's capacity to solve such collective action problems is critical for the security, sustainability and development of its members' welfare. This being the case, it is surprising how little statistically-based empirical evidence we have on how communities in developing countries solve the social dilemmas they face, about the sanctioning systems they put in place, and their effectiveness in promoting cooperation.

The literature on CPR management is rich in qualitative evidence relating to the importance of sanctions (e.g., McKean (1992), Wade (1988), Ostrom (1990), Tang (1992, 1994), Baland and Platteau (1996)). However, the case study approach generally applied in this field renders it difficult to isolate the effects of the various factors supporting cooperation and, thus, prevents the drawing of conclusions about their relative importance. In addition, while the approach successfully identifies critical factors like the need for sanctions, it does not tell us how, *ceteris paribus*, a community that is less willing or able to sanction uncooperative behaviour performs relative to a community that is more willing or able to sanction when faced with a social dilemma. The few quantitative studies of CPR management have tended to focus on other issues such as the importance of individual incentives, leadership, and group heterogeneity (e.g., Gaspart, Jabbar, Melard, and Platteau (1996)), (Baland and Platteau (1999)). Platteau and Seki (1999) and Gaspart and Seki (2000) draw conclusions about the role of status and related social sanctions in supporting collective action from an analysis of individual boat performance within three traditional Japanese fishing communities. However, these conclusions are based on indirect evidence relating to the distributions of catch size within the communities and the extent to which they reflect the distributions of crew capabilities. The qualitative data they have on the perceived importance of status and the role of social sanctions is used only to distinguish the three communities one from another and to direct the interpretation of the findings from the performance analysis. Essentially they only have three observations and so, while the analysis is both fascinating and carefully executed, it does not constitute a formal statistical test of the efficacy of social sanctions. The greatest body of quantitative empirical work on social dilemmas relates to risk sharing arrangements, although here the exclusive focus has been the extent to which and not how communities have solved the dilemma and thereby moved towards the social optimum (Udry (1994), Townsend (1995), Ravallion and Chaudhuri (1997), Grimard (1997), Jalan and Ravallion (1999)). Several theoretical models have been developed that explore the institutional conditions under which cooperation can be sustained (Coate and Ravallion (1993), Platteau (1997)), but the efficacy of different monitoring and sanctioning arrangements remains to be explored empirically.

One reason for the lack of statistically based research into the efficacy of different sanctioning systems relates to the problem of identifying samples of similar cooperative potentials. Another relates to how to measure degrees of cooperation both at the individual and the collective level. And a third to how to quantify and measure the sanctions that are being applied or, even more elusively, anticipated. Here, I solve the first of these problems by providing a sample of communities in sub-Saharan Africa with the same simple opportunity for collective action. The second and third problems can then be solved by carefully controlling the environment in which the opportunity is presented to ensure that the behaviour of interest can be observed and recorded. The opportunity

presented to the communities took the form of a public goods (PG) game played for real money (between a half and two day's casual labour wage) that the community members got to keep and played under conditions similar to those applied in experimental economics laboratories throughout Europe and the US.

Ten village communities were invited to play a first set of PG games, which I henceforth refer to as the first experiment. In each community the PG game was played twice under conditions that prohibited or at least made very difficult the imposition of sanctions: each villager chose his or her level of cooperation in private and no information linking individual villagers' identities to their levels of cooperation was circulated. Then, in five of the communities a third game was played under conditions that allowed the villagers to impose sanctions and under which they might have reasonably anticipated sanctions had they behaved in a way that offended their neighbours. Here, the villagers knew that having decided on their level of cooperation they would have to announce it to everyone present. With full information, the villagers could have imposed sanctions in whatever form they saw fit after the experimental session had ended and each of them would have known this when making their decision. In the remaining five villages the third game was played under precisely the same conditions as the first and second, thereby providing a control.

Discussions with key informants from two of the communities that took part in this experiment revealed that the chosen form of sanctioning being applied after experimental sessions was light-hearted criticism or the withholding of praise during informal gatherings. This prompted and guided the design of a second experiment conducted in thirteen communities none of whom had participated in the first experiment. Here, a discussion round following the announcements in the third game provided the villagers with an opportunity to sanction through criticism or the withholding of praise while still under observation. Then a fourth game was played, during which levels of cooperation were, once again, announced.

This experimental approach has its drawbacks. We must be mindful of the abstract nature of the collective opportunity; its relative economic unimportance compared to the everyday social dilemmas the communities face; its short time frame and finite nature; the contrived conditions under which the decisions are made; the need for any sanctions applied to emerge instantaneously; and the potential impact of observation by outsiders. All these factors make the opportunity presented to the communities very unusual and require that caution be applied when generalizing the resulting conclusions to naturally occurring dilemmas. This notwithstanding the approach yields data that can be used to explore several issues relating to cooperation and sanctioning behaviour that would otherwise be beyond our scope. Here, using the data from the first experiment, I look at whether, when given an opportunity to sanction, communities raise their levels of cooperation; and whether there is variation across communities in this regard. Using the data from the second experiment I look at which decisions lead to the imposition of shame-based sanctions by others and the effect that those sanctions have on subsequent behaviour. I also look at what the results imply about the preferences of sanctioners and sanctionees. The main empirical findings are: that out of the eighteen communities in which the third game was played with full information (five in the first experiment and thirteen in the second), fourteen significantly increased their levels of cooperation; that in all of the thirteen communities involved in the second experiment, low contributions attracted sanctions in the form of criticism; that in six of the thirteen communities, high contributions also attracted criticism; that there was significant variation in the communities' willingness to sanction through criticism; that those who were criticised for making low contributions significantly increased their contributions in the subsequent game; that those who were criticised for making high contributions significantly reduced their contributions in the subsequent game; and that those who were not criticised themselves but witnessed the criticism of

others who made the same contributions as they made similar, if not stronger, adjustments to their behaviour.

The paper has five sections. Following this introduction, section 2 provides the theoretical framework used when interpreting the experimental results later in the paper. It looks briefly at several theories that can explain why individuals impose sanctions and at the assumptions required to ensure that various types of sanction work. In sections 3 and 4, I present the first and second experiments respectively. Sub-sections 3.1 and 4.1 contain the details of the experimental designs, including a brief description of the communities that took part in the study. Sub-sections 3.2 and 4.2 contain the results. In section 5 I discuss the results and draw some tentative conclusions.

2. The theoretical basis for sanctioning behaviour

Like cooperation in one-shot social dilemmas, the imposition of costly sanctions cannot be justified under assumptions of selfish utility-maximization. Further, these predictions generalize to situations in which the social dilemma is finitely repeated. This notwithstanding, Ostrom, Walker, and Gardner (1992) and Fehr and Gächter (2000) show that when experimental subjects in finitely repeated CPR and PG games are allowed to impose fines on other subjects at a cost to themselves, a significant proportion do precisely that and levels of cooperation increase as a result. In an effort to explain both cooperative and sanctioning behaviour, theoretical economists have explored several motives.

First, sanctions might be emotionally triggered responses to violations of fairness norms. Under Levine's (1998), Fehr and Schmidt's (1999), or Bolton and Ockenfels' (2000) assumptions of aversion to inequality, sanctions would be imposed upon non-cooperators by cooperators because by not cooperating the former are preventing the latter from getting their fair share. If the imposition of sanctions reduces the payoff to the sanctionee more than the payoff to the sanctioner, cooperators can redress this imbalance by sanctioning non-cooperators. Fehr and Schmidt's (1999) model can also accommodate a preference for advantageous inequality, i.e., the taste for a positive difference between one's own and others' payoffs. Under these preferences, if the imposition of sanctions reduces the payoff to the sanctionee more than the payoff to the sanctioner, cooperators might also be sanctioned. Here, the sanctioners will be non-cooperators, as cooperation is inconsistent with the preference for advantageous inequality. A similar motivation for the sanctioning of cooperators by non-cooperators can be derived from Levine's (1998) model of altruism and spitefulness.

Second, sanctioning can be motivated by a desire to reciprocate. According to Rabin's (1993) and Dufwenberg and Kirchsteiger's (1999) theories of reciprocity a cooperating agent would gain utility from sanctioning an act of non-cooperation if she judged the latter as being unkind and intended to do her harm and the former as being an appropriately unkind retaliatory act. If we introduce the emotion of shame, these theories of reciprocal action can also explain the sanctioning of cooperators.¹ If a non-cooperator perceives an act of cooperation by another as an effort to shame him, then he will retaliate by sanctioning the cooperator. The model of social status proposed by Jones (1984) provides a similar motivation for the sanctioning of cooperators. Here, cooperation

¹ According to Scheff (1990) shame arises 'out of the monitoring of one's own actions by viewing one's self from the standpoint of others' (p. 80). Williams (1993) concurs, stating that it is associated with the experience of being seen behaving inappropriately by either an internal or an external witness.

generates a negative externality for those who want to neither cooperate nor suffer a loss of social status because they are seen to be uncooperative.²

Once we move to infinitely repeated games or introduce discounting or uncertainty into finitely repeated games, sanctioning becomes consistent with selfish utility-maximization (Kreps, Milgrom, Roberts, and Wilson (1982), Fudenberg and Maskin (1986)). Under such strategic sanctioning, cooperators are prepared to sanction non-cooperators in order to get them to cooperate in subsequent games, although, in equilibrium the threat of sanctions deters non-cooperation and no actual sanctioning is necessary. There is no strategic motivation for the sanctioning of cooperators.

Falk, Fehr and Fischbacher (2000) endeavour to establish which of these theories best explains sanctioning behaviour. They conducted a series of experiments designed to distinguish between them. They found no evidence of strategic sanctioning and that models based on reciprocity explained a greater proportion of the observed behaviour than models based on inequality aversion. In addition, they observed sanctioning of cooperators by non-cooperators, which they attributed to a preference towards advantageous inequality. Although at least some of this behaviour could also have been motivated by reciprocity combined with internal shame.

There is another condition that must be met before individuals will impose sanctions: the sanctions have to be effective, i.e., they must lead to a decline in the utility of the sanctionee. To establish how different types of sanction do this, we need to shift our focus away from the preferences of sanctioners towards those of sanctionees. Consider, first, the sanctions imposed in the experiments of Ostrom, Walker, and Gardner (1992), Fehr and Gächter (2000), and Falk, Fehr and Fischbacher (2000). These were costly fines that reduced the sanctionee's monetary payoff. Selfish utility-maximizing sanctionees would have suffered a loss in utility as a result. No other regarding preferences are required to make this type of sanction effective. The same is true of sanctions involving exclusion, imprisonment, the destruction or removal of property or any other act that leads to a loss in earnings or wealth.

In contrast, non-pecuniary sanctions such as criticism, ritual humiliation, and ostracism are ineffective under the assumption of selfish utility maximisation. Here, in order for a sanctionee to suffer a loss in utility, he must care about what others think, i.e., he must feel external shame.³ To the extent that a sanctionee can imagine what others are thinking, no action except observation is required on the part of the sanctioners. However, any manifestation of disapproval on the part of the sanctioners will remove any doubt about what they think and may also draw others' attention to the sanctionee and thereby deepen the shame. Kandel and Lazear (1992) develop a model that explicitly takes account of such shame-based sanctions or 'peer pressure' as they prefer to call it. An individual's utility is affected by her own and others' levels of cooperation and additional actions taken by others that do not directly affect the outcome of the collective action. These additional actions include the act of observing and manifestations of disapproval. Individuals who feel external shame respond to anticipated shame-based sanctions just as they respond to anticipated pecuniary sanctions; they choose a level of cooperation that equates the marginal expected loss in utility due to feeling external shame with the marginal loss in utility due to cooperating. Kandel and Lazear (1992) show that higher levels of cooperation but not necessarily higher levels of utility result if individuals are sensitive to peer pressure.

² Anything that motivates a cooperator to sanction a non-cooperator can also explain conditionally cooperative behaviour.

³ Recall that internal shame is felt even if no one observes or manifests disapproval. Like guilt, it acts as an internal sanctioning device (Posner and Rasmusen (1999)) that, if effective, would render sanctioning by others redundant. Thus, internal shame can be modelled as a taste for fairness or inequality aversion.

Given the particular form that preferences must take if manifested disapproval is to be effective as a sanctioning device, their prevalence in the case study literature on CPRs is worthy of note. Real world examples of such shame-based sanctions include forced public apologies, ritual humiliation occasionally involving some form of labelling, and partial and complete ostracism. More extreme shame-based sanctions were reserved for more serious violations of CPR management rules. One reason for the frequency with which such non-pecuniary sanctions are observed may be their relatively low cost. The sanctioner will incur a cost associated with the effort of observing and then manifesting her disapproval. If inappropriate manifestations of disapproval are sanctionable and if there is some uncertainty about whether her planned sanction is appropriate, there will be an additional stochastic component to her private cost function.⁴ Further, both the sanctioner and others may feel empathy for the sanctionee and suffer a loss of utility as a result. These costs may be significant, but few or no resources apart from some labour or effort are required as an input, no one is excluded from exchange, no collective income earning opportunities or increasing returns to scale are left unexploited, and no physical assets are destroyed. In addition, with shame-based sanctions a single action can not only cause a loss in utility to the sanctionee but also inform others of the sanctionee's misconduct, thereby effecting coordinated sanctioning and greater efficiency.

The literature on naturally occurring CPRs provides no statistically-based evidence of the efficacy of non-pecuniary sanctions. However, there is some experimental evidence that the anticipation of such sanctions leads to higher levels of cooperation. Gächter and Fehr (1999) designed an experiment in which some groups of subjects in finitely repeated PG games knew that they were going to discuss their own and others' levels of cooperation post-play while others knew that they were to remain anonymous throughout.⁵ Gächter and Fehr (1999) proposed that the discussion provided subjects with an opportunity to impose shame-based sanctions by criticising one another's actions. They found that the anticipated discussion had no effect on levels of cooperation when the subjects met for the first time after play at the start of the discussion. However, when they were provided with an opportunity to interact socially prior to play the anticipated discussion led to increased levels of cooperation. These results suggest that shame-based sanctions induce cooperation but only when combined with some familiarity. Further, they would lead one to predict that in communities in which individuals interact on a regular basis, shame-based sanctions will be effective. This notwithstanding, in the only previously documented economic experiment similar to Gächter and Fehr's (1999) that has been conducted within a community, Henrich and Smith (2000) found no change in levels of cooperation between an anonymous treatment and a treatment in which decision were announced making sanctioning possible. Henrich and Smith (2000) take this as evidence that the Amazonian ethnic group from which their subjects were drawn, share no concept or behavioural rules relating to equity or fairness and have no preference for punishing apparently unfair behaviour. They present ethnographic evidence that further supports this conclusion.⁶ The efficacy of shame-based sanctions within real world communities cannot be taken as given.

⁴ One would expect inappropriate shame-based sanctioning to be itself sanctionable when individuals have something to gain from slandering others (Baland and Platteau (1996)).

⁵ Gächter and Fehr (1999) made no attempt to quantify the discussion and incorporate it into their analysis.

⁶ The Machugenga display a similar disinterest in fairness when playing ultimatum games (Henrich (2000)).

3. The impact of an opportunity to sanction on levels of cooperation

3.1 Experimental design

Both experiments were conducted in rural Zimbabwe. Of the 10 communities involved in the first experiment, nine were resettled shortly after independence in the early 1980s.⁷ These resettled villages are distributed across two ‘resettlement areas’ both to the southwest of the capital, Harare. The one non-resettled village is located on the boarder of one of these areas. Resettled villages are socially distinct from non-resettled ones: compared to a non-resettled household a resettled one is significantly less likely to be related to another household in the same village. Two decades ago the resettled villages were no more than collections of strangers removed from the inherited social institutions of the villages in which they were born. Consciously or subconsciously these collections of strangers would have had to develop their own social institutions and approaches to social dilemmas. Anecdotal evidence suggests that they have succeeded in this endeavour to varying degrees. This made them an attractive focus for my study as variation facilitates the identification of causal linkages. This notwithstanding it is important to remember that the sample may not be representative of Zimbabwe as a whole. Wherever possible I test for variations in behaviour between resettled and non-resettled communities.

In seven of the communities included in the first experiment a random sample of households was invited to send one person over the age of fourteen to the experimental session.⁸ In the remaining three villages every household was invited. Any nominee who was unable to identify and rank three numbers was asked to return home and find another nominee to take their place. A total of 248 nominees took part in the first experiment. Eight of the sessions involved twenty nominees each. The remaining two involved 38 and 50 nominees. Each of the nominees participated in the experimental session that took place in their own village. The experiments were conducted outside with the nominees seated approximately two metres apart. They were asked not to talk or attempt to communicate in any way with one another and were constantly watched. They were taught the game by a shona-speaking research assistant and were then asked a series of questions before proceeding to play the games. The description of the games, the questions, and the nominees’ instructions were all scripted and the script was adhered to in all ten sessions. Nominees’ questions were answered by referring to the script. The research team (one researcher and four Zimbabwean assistants) was the same for every session and efforts were made to standardise actions and demeanours.

In the PG game each player is placed in a group containing n members (themselves and $n-1$ others) and given an initial endowment, y . Each then has to decide how much of this endowment to contribute to a public good, g_i ($0 \leq g_i \leq y$). Decisions about g_i are made simultaneously. The sum of the group’s contributions is multiplied by a factor a ($1 < a < n$), and the resulting amount is shared equally among the n nominees. Thus, the final payoff to each player is $\pi_i = y - g_i + \frac{a}{n} \sum_{j=1}^n g_j$. For an anonymous, selfish money-maximiser the dominant strategy in this game is to free-ride by setting g_i equal to zero. This is because $\partial \pi_i / \partial g_i = -1 + a/n < 0$. However, it follows from $\frac{\partial}{\partial g_i} \sum_{i=1}^n \pi_i = -1 + a > 0$ that the group’s payoff, $\sum_{i=1}^n \pi_i$, is maximized if each member contributes all of their endowment, i.e., sets g_i equal to y . In Zimbabwe the nominees’ initial endowments, y , were set equal to

⁷ Five additional communities were involved in the piloting of the first experiment. During the pilot I made significant changes to the ways in which we selected the nominees and presented the game. To eliminate the possibility of these changes biasing the reported results, I have not used the data from the pilot in the analysis presented below.

⁸ The sampling proportions varied between 0.31 and 0.59.

Zim\$100.⁹ The number of nominees per group, n , was set to five and the factor, a , was set to two. Thus, the marginal payoff to the public good was 0.4. The games were played with pens and specially designed forms. The forms presented the nominees with a set of six contribution levels (Zim\$0, Zim\$20, Zim\$40, Zim\$60, Zim\$80, and Zim\$100). They had to select one of these by drawing a circle around it. Assistance was provided to those who had difficulty drawing their circles, but they had to make their own decision about the level of contribution. After the selections had been made, the forms were collected and the calculations executed. Both the nominees' shares from their group's public good, $\frac{a}{n} \sum_{j=1}^n g_j$, and their final payoffs, π_i , were written on the forms before returning them to the nominees.

In the first experiment the nominees played three games. At the outset they were told that they would be playing several games, but were told neither the exact number nor whether the games would be identical or different.¹⁰ They were also told that at the end of the session one of the games would be picked at random (by drawing a number from a hat) to determine their earnings. The same groups of five nominees were maintained throughout the sessions and the nominees were reminded of this at the start of the second and third games.¹¹ Each player knew that the other four nominees in their group were both from their village and present in the session. However, they did not know their exact identity. Thus, their initial expectations about other nominees' behaviour would have been formed on the basis of knowledge gleaned during everyday life. In the one session where the number of nominees was not a multiple of five, one of the groups contained two virtual players. These virtual players always contributed the village mode in the current. The virtual players' contributions are not included in the analysis but those of the three nominees who were grouped with them are included.

In five of the ten villages the game was played twice anonymously and then once publicly. In the anonymous treatment each nominee selected their level of contribution in private. In the public treatment, prior to selecting their level of contribution, the nominees were told that once they had made their selections on their forms they would be asked one-by-one to announce their level of contribution to everyone present in the session. In the other five villages all the games were played anonymously, thus providing an experimental control. (See Figure 1) This first experiment closely resembles Gächter and Fehr's (1999) treatment with pre-play social interaction.

3.2 Results

Here, my primary objective is to establish whether the move from the anonymous to the public decision-making environment is associated with a change in contributing behaviour. However, first it is useful to take a look at the contributions made by players in the first game. Table 1 shows that the mean contribution for the full sample of 248 nominees in the first game was Zim\$50.65, which

⁹ Zim\$100 was approximately equivalent to US\$2 at the time of the experiments.

¹⁰ To an experimental economist, this will appear odd. Normally they would state the number of games to be played so as to elicit end game effects. However, in addition, they would go to great lengths to ensure that their subjects did not know one another or, at least, did not know that they knew one another. That my nominees knew one another is a critical feature of my study. Thus, to assume that the last game that they played in my experimental sessions was an end game could be erroneous under any treatment and would be grossly misleading under one of the treatments that I applied. By choosing to be ambiguous I guarded against such an assumption ever being made.

¹¹ By preference, experimental economists would conduct experiments based on PG games twice, once with stable groups across games, as I have done, and once with the remixing of groups between games. The former is referred to as the 'partner treatment' and the latter as the 'stranger treatment'. I chose to do only the partner treatment for two reasons, to guard against the word 'stranger' being associated with the data, and to maximize the chance of the Zimbabwean nominees learning and behaving strategically.

is within the range that Ledyard (1995) describes as usual in PG games (40 to 60 percent of the initial endowment). This notwithstanding, the means for the communities span a wide range, with two falling below 40 percent and one falling above 60 percent. A comparison the mean contributions by game for those nominees facing the anonymous treatment in all three games and those facing the anonymous treatment in the first and second game and the public treatment in the third game, reveals that there are significant differences between the two groups even before the treatments diverge: those facing the anonymous treatment throughout contributed significantly less in the first game.¹² This difference could be indicative of an inherent variation in preferences between the two groups of communities. If this is the case, a cross-section analysis of the contributions made in the third game could generate biased results relating to the differential effects of the treatments. This can be avoided by exploiting the panel-nature of the data. Thus, in all the regressions presented in Table 2 I control for nominee, community, and group fixed effects. Observations of the dependent variable are drawn from the second and third games only.¹³ This allowed me to control for strategic behaviour using the lagged mean contributions of the other four nominees in each nominee's group along with its square, cube, and quadratic term.¹⁴ The variables of primary interest in these regressions are the dummy, *t3anon*, which takes the value one if the contribution was made in the third game under the anonymous treatment and zero otherwise and the dummy, *t3public*, which takes the value one if the contribution was made in the third game under the public treatment and zero otherwise. In both cases the basis for comparison is the contribution made in the second game.

The regression in the first column is based on the full sample of 496 contributions (two per nominee).¹⁵ Nominees who faced the anonymous treatment in both the second and third games made no significant alteration to their behaviour between games, whereas the nominees who faced the anonymous treatment in the second game and the public treatment in the third contributed significantly (5 percent level) more in the latter. On average, they increased their contributions by \$5.2 after controlling for strategic behaviour. Columns (2) to (4) contain similar regressions based on sub-samples defined according to the size of the contributions made in the preceding game. Those who contributed \$0 or \$20 in the preceding game and experienced the change in treatment contributed Zim\$12.1 more in the third game (one percent level of significance), while those who made low contributions in the preceding game and remained anonymous in the third did not significantly alter their behaviour. Those who contributed \$80 or \$100 in the preceding game and remained anonymous contributed Zim\$22.9 less in the third game (one percent level of significance), while those who experienced the change to the public treatment continued to make high contributions in the third game. These results suggest that the change from the anonymous to the public treatment causes a change in contributing behaviour, although it is only in the regression based on the sub-sample who made the high contributions in the preceding game that the coefficients on *t3anon* and *t3public* vary significantly (one percent level) from one another.

In the regression presented in the fifth column of Table 2 the nominee's own contribution in the preceding game enters the model as an additional control variable. This lagged dependent variable

¹² Non-resettled nominees' behaviour is indistinguishable from resettled nominees facing the same treatment.

¹³ Recall from the experimental design that individual contributions could take one of six values. This notwithstanding, here I treat contributions as a continuous variable. When large numbers of fixed effects are introduced into multinomial probit or logit regressions they can lead to significant biases in the estimated standard errors. The inclusion of LDVs can cause similar problems. The implications of treating contributions as continuous are (a) that the regressions could predict contributions outside the range zero to one hundred and (b) that non-linear underlying relationships may not be captured. None of the regressions presented predict contributions outside the range zero to one hundred, and I have explored potential non-linearities either by including interaction terms or by splitting the sample.

¹⁴ Increasing the order of the polynomial beyond four did not improve the fit of the model.

¹⁵ Three of the 148 nominees (2 percent) announced contributions greater than their actual contributions in the third game. Throughout the analysis I use actual contributions as the dependent variable.

(LDV) may be correlated with the fixed effects and this could lead to biased coefficients (Nickell (1981)). To control for such a bias, I employ differenced generalized method of moments (GMM) estimators (Arellano and Bond (1991)). The consequent coefficient on the LDV is not significant, although its inclusion does have implications for other results of interest.¹⁶ The coefficient on *t3public* in this regression is conceptually equivalent to the difference between the coefficients on *t3anon* and *t3public* in columns (1) to (4). This coefficient is significant (5 percent level). A switch to the public treatment in the third game is associated with an inter-game increase in contributions that is Zim\$5.3 greater than any increase made when anonymity persists.

Finally, treating the regression in Column (1) of Table 2 as the basic model, I tested whether each of the communities responded differently to the opportunity to impose sanctions. I interacted the *t3public* dummy with four village dummies, each corresponding to one of the communities in which the third game was played publicly (the fifth community was the basis for comparison). The p-value corresponding to the joint test of significance for these four dummies was 0.469 suggesting that the responses of the five communities to the change in treatment were indistinguishable.¹⁷

These results concur with those of Gächter and Fehr (1999). The opportunity to sanction combined with the familiarity that undoubtedly exists within the Zimbabwean communities led to higher levels of cooperation. However, there is an important distinction between Gächter and Fehr's (1999) finding and mine: while they could be fairly sure that the increase in cooperation in anticipation of a discussion was due to fear of shame-based sanctioning, I cannot. In their experiment, once the discussion was over the participants left the lab one-by-one and could be assumed unlikely to interact in the future. In contrast, for my nominees, the game represented just another interaction in a series of uncertain length. Thus, even though it seems unlikely given the parameters of the game, we cannot rule out the possibility that the threat of subsequent exclusion from economically beneficial interactions could have motivated the increase in cooperation. This is why the stories told by key informants from the two villages were so important. They suggested that the sanctions of choice in this context were indeed shame-based. My aim with the second experiment was to draw these shame-based sanctions into the observable domain and thereby calibrate their effect. Finding no way to calibrate the potential effect of threatened exclusion, I must be content with a partial analysis.

4. Shame-based sanctioning and its ex-post impact on cooperation

4.1 Experimental Design

The second experiment involved thirteen communities. Motivated by the variation in cooperative behaviour in the first experiment, I decided to invite a larger number of non-resettled villages to participate in the second. I included four non-resettled communities and nine resettled ones. The nine resettled communities are all located in a resettlement area to the north-west of Harare. Two of the non-resettled villages are located on the border of this area and the other two are located on the border of one of the resettlement areas visited during the first experiment. In the nine resettled communities every household was invited to send a nominee to the experimental session in their village. In the non-resettled communities a random sample of 25 households were invited.¹⁸ A total

¹⁶ The results of a Sargan test of over-identifying restrictions indicates that the instruments used for the LDV in the GMM estimation are valid.

¹⁷ Similarly, there is no evidence that the nominees in the one non-resettled community respond differently to the change in treatment between the second and third game.

¹⁸ In non-resettled communities, households divide as sons marry and the number of households is often uncertain. This is not the case in resettled communities, in which the numbers of households is a matter of policy. Inviting one member

of 354 nominees were involved in the game with each community-specific session involving between thirteen and 46 nominees. So, once again, some virtual players were required to make up numbers.

In each session, four public goods games were played. The first two games were played anonymously. The third game was played publicly and was followed by a ‘discussion round’. During this round, the nominees were invited to make public verbal statements about each other’s decisions. Special care was taken not to lead the nominees. The aim was to provide them with an opportunity to comment, while leaving them free to complement, criticise, or remain silent as they saw fit. One research assistant would stand beside each player in turn and say ‘Player number ..., Mr/Mrs ..., contributed \$... Does anyone have anything to say about that?’ Two other research assistants recorded which nominees were criticised, which were complemented and which received neither criticism nor complement, and a fourth research assistant recorded as much of the substance of the criticism as time allowed. These three independent records were reconciled directly after the experimental session and the analysis presented below is based on the reconciled data.¹⁹ Once the discussion round was complete, a fourth PG game was played. Like the third, this game was played publicly. The schematic for the second experiment is depicted Figure 2.

As before, the nominees were told at the start of the session that they would be playing several games, but were told neither the exact number nor whether the games would be identical or different. They were told that at the end of the session one of the games would be picked at random (by drawing a number from a hat) to determine their earnings, and were repeatedly reminded that the same groups of five nominees were to be maintained throughout. Each player knew that the other four nominees in their group were both from their village and present in the same session as they, but did not know their exact identity.

4.2 Results

Table 4 contains summary statistics relating to the patterns of criticism and praise that emerged during the discussion rounds. Of the 354 nominees that participated in this experiment 29.4 percent were criticised, 24.9 percent for contributing too little and 4.5 percent for contributing too much, 65.3 percent were praised, and only 5.4 percent received no comment. The pattern of criticism and praise varies depending on the level of contribution made by the nominee being discussed. Those who contributed \$0 or \$20 were most likely to be criticised for contributing too little (67.7 percent), while those who contributed \$80 or \$100 were never criticised in this way. Only those who contributed \$80 or \$100 were criticised for contributing too much, although only fourteen percent experienced such comments. More generally, these nominees were praised (83.3 percent). It is useful to know a little more about the nature of the criticism that was observed. Those who were criticised for making low contributions were told that they were ‘stingy’ or ‘mean’, while those who were criticised for making high contributions were told that they were ‘stupid’ or ‘careless with their money’. All of the criticism was delivered in a light-hearted manner and was usually accompanied by a great deal of laughter on the part of the critics, the criticised, and the other nominees. Finally, at no point did the criticism appear to contain information that might harm the

from every household in a resettled community yields a predictable number of nominees, whereas in a non-resettled community it yields an uncertain and very large number of people. This is why I applied different sampling procedures for the two types of community.

¹⁹ We were aware that, even though it was designed to mimic a process that was naturally occurring in at least some of the villages, the discussion round could cause offence. For this reason we held group discussions after each session and, wherever possible, also followed-up with trusted key informants. Neither of these follow-up approaches yielded any indication that the villagers were unhappy with the games.

reputation of the player independently of the harm caused by his or her announcement of the contribution he or she made. The criticisms that got closest to containing reputation-relevant information took the form of jokes like ‘Now I know why I never get offered food when I drop by your house!’ In such instances the response of the other villagers suggested that what they were hearing was either old news or not to be taken seriously.

The summary statistics in Table 4 suggest that the likelihood of criticism is non-linearly related to contributions. This non-linearity is also clear from the graphs presented in Figures 3 and 4. In each graph in Figure 4 the proportion of nominees criticised in the discussion round is plotted against the contributions made in the third game. There is one graph for each of the thirteen communities in which there was a discussion round. Five the graphs are somewhat U-shaped: both high and low contributors are more likely to be criticised than those contributing a moderate amount. This U-shape is also evident in Figure 3, which depicts the likelihood of criticism for each level of contribution of the full sample of 354 nominees, and is found to be statistically significant in the Probit analysis presented in Table 5. Here, *criticised*, a dummy variable that takes the value of one for a nominee who is criticised for making either too high or too low a contribution and zero otherwise, is the dependent variable and the coefficient on the square of the contribution in the third game is positive and significant at the one percent level.

Both the graphs and the regressions also highlight the considerable variation in propensities to criticise across communities. Compare the graphs in the top row of Figure 3 with the first three in the second row. In all but one of the twelve possible combinations, the graph from the top row dominates (is above and to the right of) the graph from the second row. From the probit analysis we see that after controlling for contributions, the likelihood of being criticised varies over a range of 60 percentage points across the communities and that the community dummies are jointly highly significant (one percent level).²⁰

Table 5 presents the mean contributions by game for the full sample of 354 nominees. The levels of contribution are similar to those made by the subjects who played the three anonymous games in the first experiment. There are only marginal changes in mean contributions between the second, third, and fourth games. The spread in the community means considerable in every game and the variation in community means is significant in the first, third, and fourth games. Finally, the spread is greatest and the variation most significant in the fourth, post discussion, game.²¹

Table 6 contains a regression analysis of contributions that takes account of the effects of criticism on subsequent contributing behaviour. Due to the significant between-community variations in contributions, I once again exploit the panel-nature of the data to control for nominee, group and community fixed effects. As before, lags of the mean contributions made by other members of the nominees’ groups and powers of the same are used to control for strategic behaviour. The effects of changes in treatment are captured by three dummy variables, *public*, *post-crit*, and *public-ns*. The first and second of these identify the third game (public treatment) and the fourth game (public, post-discussion treatment) respectively. The third dummy, *public-ns*, identifies contributions made in the third game in four communities where the change to the public treatment was associated with a significant reduction in contributions. Currently, no explanation can be give as to why these four communities responded differently. The significance of the positive coefficient on *public* indicates that in the remaining nine communities the change in treatment had the expected positive effect on contributions. In two out of the three regressions presented the negative coefficient on *post-crit* is

²⁰ Replacing the set of community dummies with one dummy distinguishing resettled from non-resettled nominees reveals that former are significantly less likely to criticise (10 percent level). After controlling for contribution level, the likelihood of being criticised is twelve percent lower in resettled communities.

²¹ Mean contributions by resettled nominees were marginally but significantly higher (5 percent level).

significant suggesting that, *ceteris paribus*, the nominees reduce their contributions in the post-discussion game.

The regressions also contain two dummy variables, *criticised-low* and *criticised-high*, which capture the individual nominees' experiences in the discussion round. *Criticised-low* is a dummy variable that takes the value one if the contribution was made after the nominee had been criticised for making too low a contribution in the third game. *Criticised-high* is a dummy variable that takes the value one if the contribution was made after the nominee had been criticised for making too high a contribution in the third game. The regression in column (1) suggests that those who were criticised for making low contributions contributed significantly more (one percent level), while those who were criticised for making high contributions contributed significantly less (five percent level) in the subsequent game.²²

However, these results do not take account of the possibility that criticism has an effect not only on those who experience it first hand, but also on those who are witnesses. I explore these effects in column (2) by introducing two additional variables, *prop-crit-low*, which equals the proportion of nominees who were criticised for making a low contribution having made the same contribution as nominee *i* in the preceding game and in the same community, and *prop-crit-high*, which equals the proportion of nominees who were criticised for making a high contribution having made the same contribution as nominee *i* in the preceding game and in the same community. The coefficients on both of these variables are large, significant and bear the expected sign. The more criticism of low contributors a low contributor witnesses, the greater the upward adjustment he makes to his subsequent contributions. And the more criticism of high contributors a high contributor witnesses, the greater the downward adjustment he makes to his subsequent contributions. In addition, the inclusion of these variables in the model renders the coefficient on *criticised-high* insignificant and changes the sign but not the significance of the coefficient on *criticised-low*. These results suggest that it is the overall pattern of criticism within a community rather than their individual experience that affects nominees' subsequent behaviour. Further, those who are criticised for making low contributions increase their contributions by significantly less than those who just witness the criticism of others. These results are strengthened by the introduction of a LDV into the model (see column (4) in which GMM estimators are presented). The coefficient on the LDV is positive and significant (one percent level) suggesting that after controlling for other factors the nominees' tend to over-correct or overreact to the stimuli (evidence of what others in their group have done and the discussion) they receive.

The coefficients on *prop-crit-low*, *prop-crit-high*, and *criticised-low* are most readily understood through simulations. Figure 5 contains three simulations based on the coefficients relating to these three variables from column (4) of Table 6. Consider a community in which nominees' contributions in the third game are uniformly distributed as in the first row of Figure 5 and as a consequence the mean contribution is Zim\$50. Now, holding everything else constant and taking account of the effects of witnessing criticism only, consider what happens if the community's incidence of criticism curve as that depicted in the second row and column (a); only twenty percent of those making contributions of Zim\$0 and Zim\$20 are criticised. Those who made low contributions and witnessed low contributors being criticised will increase their contributions in the fourth game yielding the distribution presented in the third row of column (a). If there is more criticism of low contributors, as suggested by the curve in column (b), the impact on witnesses' contributions in the third game is greater and the distribution in the third row of column (b) results. But if there is also criticism of high contributors as in column (c), then as well as low contributors adjusting upwards, high contributors will adjust downwards resulting in the distribution depicted in

²² Coefficients on salient interaction terms suggest that resettled and non-resettled nominees respond in the same way to the change to the public treatment and to the discussion round.

the third row of column (c). In the fourth row I take account of the effect of first-hand experience of being criticised for making a low contribution. In column (a) few people experienced such criticism first hand so the effect is negligible. However, in columns (b) and (c) the high incidence of people who experienced such criticism first hand leads to a considerable reduction in the extent to which the graph shifts to the right. This notwithstanding, even after taking account of this first-hand effect, more criticism of low contributors leads to a greater rightwards shift in the distribution and a higher mean contribution.

5. Discussion and conclusions

These results provide strong evidence that the shame-based sanctions anticipated and imposed by the communities that took part in my experiments were effective at promoting cooperation. Villagers in Zimbabwe clearly care about what other people think of them and will modify their behaviour in order to improve their status in the eyes of their neighbours. This being the case, shame-based sanctions could be used strategically. Indeed, given that these games could have been viewed as additions to the life-long series of interactions between neighbours, we cannot rule out the possibility that some of the nominees were sanctioning strategically. However, those who sanctioned high contributors can only have been motivated by shame, by a desire either to punish someone who has made them (or others, if they also feel empathy) look bad or to encourage someone to behave in a way that makes them (or others) look less bad in the future. The first of these explanations is consistent with reciprocity in the presence of shame, while the second is consistent with status maximization.

The results also show that the effect of shame-based sanctions on cooperation is far from straightforward. In particular, it is the overall pattern of shame-based sanctioning rather than a nominee's individual experience that has a corrective impact on behaviour. This could be because the actual criticism in the discussion round allowed the nominees to build a picture of what their neighbours considered acceptable. We could be observing a learning process and convergence on an equilibrium in which nominees contribute in accordance with their obligations to their communities and obligations are defined with reference to the level of contribution that each nominee would like all community members to choose (Sugden (1984)). Alternatively the nominees might have viewed the criticism of others who behaved in the same way them as being directed towards them also.

The analysis identifies four distinct effects of shame-based sanctions. First, those who make low contributions and witness the criticism of others who have made similar contributions, subsequently make higher contributions. So, making an example of someone appears to work. Second, those who experience criticism directly make less of an upward adjustment to their contributions than those who are merely witnesses. This could be an endogenous effect: the communities may have known which individuals needed the most encouragement to increase their levels of cooperation; they may have reserved direct criticism primarily for those who were shameless. Alternatively we could have been observing a retaliatory action: the criticised continued to behave less cooperatively because they wished to punish those who criticised them. Third, those who make high contributions and witness the criticism of others who have made similar contributions, subsequently make lower contributions. So, individuals can be shamed into behaving more selfishly by others who are trying to save face. Finally, returning to the significant negative coefficient on *post-crit*, individuals in communities that make little or no use of opportunities to impose shame-based sanctions subsequently cooperate less. The absence of criticism seems to work as a vindication of more selfish behaviour. The first of these effects is the most powerful and is also the only one that facilitates cooperation. Each of the others diminishes cooperation either directly or by lessening the impact of the first, pro-cooperative effect.

The analysis is not without limitations. Of greatest concern are the extremely low R^2 s reported in the tables. Either the data is very noisy or there are some important variables being omitted from the analysis. That there is considerable noise in the data is undoubtedly true. With few exceptions, the mathematical abilities of the Zimbabwean nominees were poor, especially compared to those of the OECD undergraduates who are usually selected to participate in such games. But this does not preclude omitted variable bias. The fixed effects will have fully controlled for all unobserved nominee, community, and group characteristics. However, they will not have controlled for differences in responsiveness to changes in treatment and experience across nominees. The four communities where cooperation fell following the change to the public treatment are cases in point. Future research might usefully investigate this and other differential impacts. In particular, it would be interesting to look at how the characteristics of those giving and receiving criticism affect its efficacy as a sanctioning device.

What can we conclude about the experimental methodology applied in this study? The social dilemma presented to the participating communities was unarguably alien to them. Yet during the course of the experiments they demonstrated, first, that they held views about how they should each respond to the dilemma; second, that they had views about how their neighbours should respond; third, that they could readily adapt the shame-based sanctioning system that they presumably apply in their everyday lives to the new dilemma; and fourth, that they could effect greater cooperation or different patterns of cooperation as a result. If they can do this in the context of an alien social dilemma in only a couple of hours and with only one opportunity for bilateral verbal interaction, it seems reasonable to assume that they can also do it when faced with more familiar social dilemmas. A more contentious question is whether the variations in cooperative and sanctioning behaviour observed during the experiments reflect unobserved variations in behaviour in real social dilemmas. Future research will aim to answer this question by comparing the experimental results with data on real social dilemmas. In the mean time we must be satisfied to conclude that at least in the context of this social dilemma, in those communities where individuals are more willing to criticise non-cooperators and less interested in criticising cooperators in order to save face, more cooperative outcomes result.

References

- Arellano, M. and S. Bond (1991) 'Some Tests of Specification for Panel Data: Monte Carlo Evidence and Application to Employment Equations,' *Review of Economic Studies*, 58, pp. 277-297.
- Baland, J. M. and J. P. Platteau (1996) *Halting the Degradation of Natural Resources – Is there a Role for Rural Communities?* Oxford: Clarendon Press.
- Baland, J. M. and J. P. Platteau (1999) 'The Ambiguous Impact of Inequality on Local Resource Management', *World Development*, 27, pp. 773-88.
- Bolton, G. E. and A. Ockenfels (2000) 'A Theory of Equity, Reciprocity, and Competition,' *American Economic Review*, 90(1), pp. 166-93.
- Coate, S. and M. Revallion (1993) 'Reciprocity without Commitment: Characterization and Performance of Informal Insurance Arrangements,' *Journal of Development Economics*, 40(1), 1-24.

- Dufwenberg, M. and G. Kirchsteiger (2000) 'Reciprocity and Wage Undercutting,' *European Economic Review*, 44, pp. 1069-78.
- Falk, A., E. Fehr, and U. Fischbacher (2000) 'Informal Sanctions,' Institute for Empirical Research in Economics, University of Zurich, Working Paper Series, ISSN 1424-0459.
- Fehr, E. and S. Gächter (2000) 'Cooperation and Punishment in Public Goods Experiments,' *American Economic Review*; 90(4), pp. 980-94.
- Fehr, E. and K. M. Schmidt (1999) 'A Theory of Fairness, Competition, and Cooperation,' *Quarterly Journal of Economics*, 114(3), pp. 817-68.
- Fudenberg, D. and E. Maskin (1986) 'The Folk Theorem in Repeated Games with Discounting or with Incomplete Information,' *Econometrica*, 54(3), pp. 533-54.
- Gächter, S. and E. Fehr (1999) 'Collective Action as a Social Exchange,' *Journal of Economic Behaviour and Organization*, 39, pp. 341-69.
- Gaspart, F., M. Jabbar, C. Melard, and J. P. Platteau (1996) 'Participation in the Construction of a Local Public Good with Indivisibilities: An Application to Watershed Development in Ethiopia,' *Journal of African Economies*, 7(2), pp. 157-84.
- Gaspart, F. and E. Seki (1999) 'Sharing, Heterogeneity and Status Considerations: Incentive Theory and Empirical Evidences,' in E. Seki *Community as a Solution to Market Failures: Pooling Groups in Japanese Fisheries*, Dissertation, Universitaires Notre-Dame de la Paix-Namur, Belgium.
- Grimrand, F. (1997) 'Household Consumption Smoothing Through Ethnic Ties: Evidence from Côte D'Ivoire,' *Journal of Development Economics*, 53, pp. 391-422.
- Henrich, J. (2000) 'Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining among the Machiguenga of the Peruvian Amazon,' *American Economic Review*, 90(4), pp. 973-79.
- Henrich, J. and N. Smith (2000) 'Comparative Experimental Evidence from Machiguenga, Mapuche, Huinca & American Populations Shows Substantial Variation Among Social Groups in Bargaining and Public Goods Behavior,' mimeo, University of Michigan Business School.
- Jalan, J. and M. Ravallion (1999) 'Are the Poor Less Well Insured? Evidence on Vulnerability to Income Risk in Rural China,' *Journal of Development Economics*, 58(1), pp. 61-81.
- Jones, S. (1984) *The Economics of Conformity*, Oxford: Basil Blackwell.
- Kandel, E. and E. P. Lazear (1992) 'Pressure and Partnerships,' *Journal of Political Economy*, 100(4), pp. 801-17.
- Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982) 'Rational Cooperation in the Finitely Repeated Prisoners' Dilemma,' *Journal of Economic Theory*, 27(2), pp. 245-52.
- Levine, D. K. (1998) 'Modeling Altruism and Spitefulness in Experiments,' *Review of Economic Dynamics*, 1, pp. 593-622.

- McKeen, M. (1992) 'Success on the Commons: A Comparative Examination of Institutions for Common Property Resource Management,' *Journal of Theoretical Politics*, 4, pp. 247-82.
- Nickell, S. J. (1981) 'Biases in Dynamic Models with Fixed Effects,' *Econometrica*, 49(6), pp. 1417-26.
- Ostrom, E. (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge: Cambridge University Press.
- Ostrom, E., R. Gardener, and J. Walker (1992) 'Covenants With and Without the Sword: Self Governance is Possible,' *American Political Science Review*, 86, 404-17.
- Posner, R. A. and E. B. Rasmusen (1999) 'Creating and Enforcing Norms, with Special Reference to Sanctions,' *International Review of Law and Economics*, 19, pp. 369-82.
- Platteau, J. P. (1997) 'Mutual Insurance as an Elusive Concept in Traditional Rural Communities,' *Journal of Development Studies*, 33(6), pp. 764-96.
- Platteau, J. P. and E. Seki (1999) 'Heterogeneity and Social Esteem in Income-Pooling Groups,' in E. Seki *Community as a Solution to Market Failures: Pooling Groups in Japanese Fisheries*, Dissertation, Universitaires Notre-Dame de la Paix-Namur, Belgium.
- Rabin, M. (1993) 'Incorporating Fairness into Game Theory and Economics,' *American Economic Review*, 83(5), pp. 1281-302.
- Ravallion, M. and S. Chaudhuri (1997) 'Risk and Insurance in Village India: Comment,' *Econometrica*, 65(1), pp. 171-84.
- Scheff, T. J. (1990) *Microsociology*, Chicago: University of Chicago Press.
- Sugden, R. (1984) 'Reciprocity: The Supply of Public Goods Through Voluntary Contributions,' *Economic Journal*, 94, 772-87.
- Tang, S. Y. (1992) *Institutions and Collective Action: Self-Governance in Irrigation*, San Francisco: CA: Institute of Contemporary Studies Press.
- Tang, S. Y. (1992) 'Institutions and Performance in Irrigation Systems,' in Ostrom, E., R. Gardener, and J. Walker, *Rules, Games, and Common-Pool Resources*, Ann Arbor: The University of Arizona Press.
- Townsend, R. M. (1995) 'Financial Structures in Northern Thai Villages,' *Quarterly Journal of Economics*, 110(4), pp. 1011-46.
- Udry, C. (1994) 'Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria,' *Review of Economic Studies*, 61, pp. 495-526.
- Wade, R. (1988) *Village Republics – Economic Conditions for Collective Action in South India*, Cambridge: Cambridge University Press.
- Williams, B. (1993) *The Moral Sense*, New York: Free Press.

Figure 1: Schema for first experiment 1

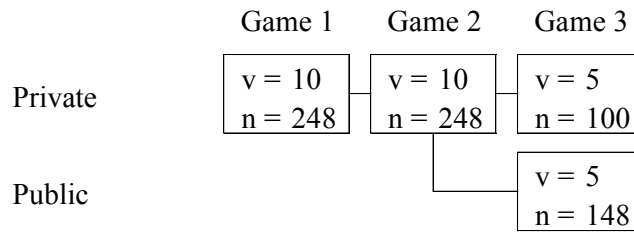


Figure 2: Schema for second experiment

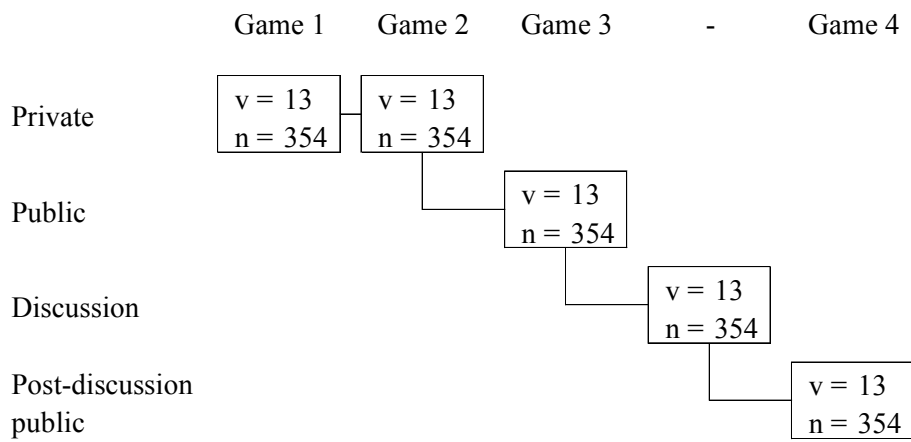


Figure 3: Incidence of criticism in full sample

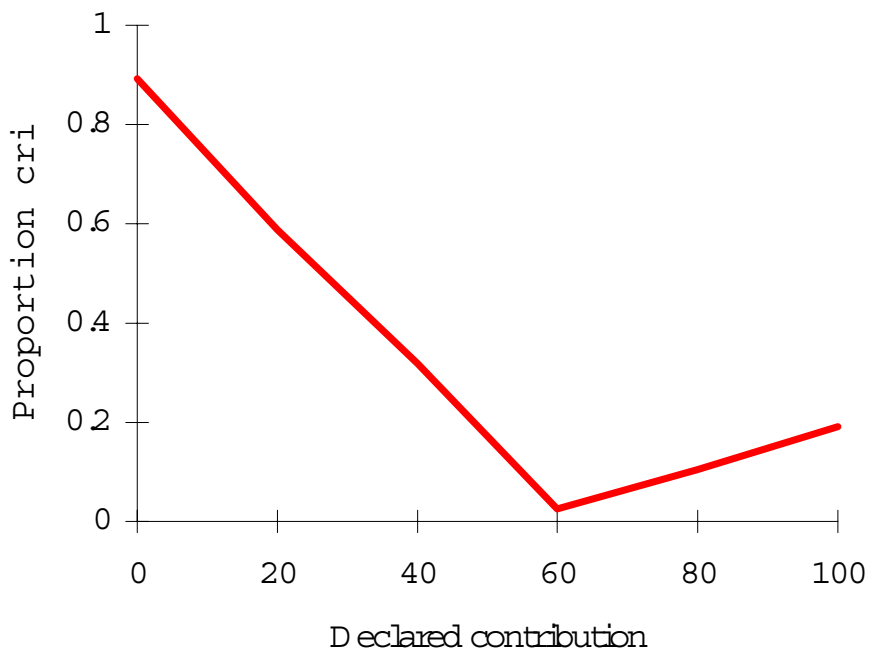


Figure 4: Incidence of criticism by village

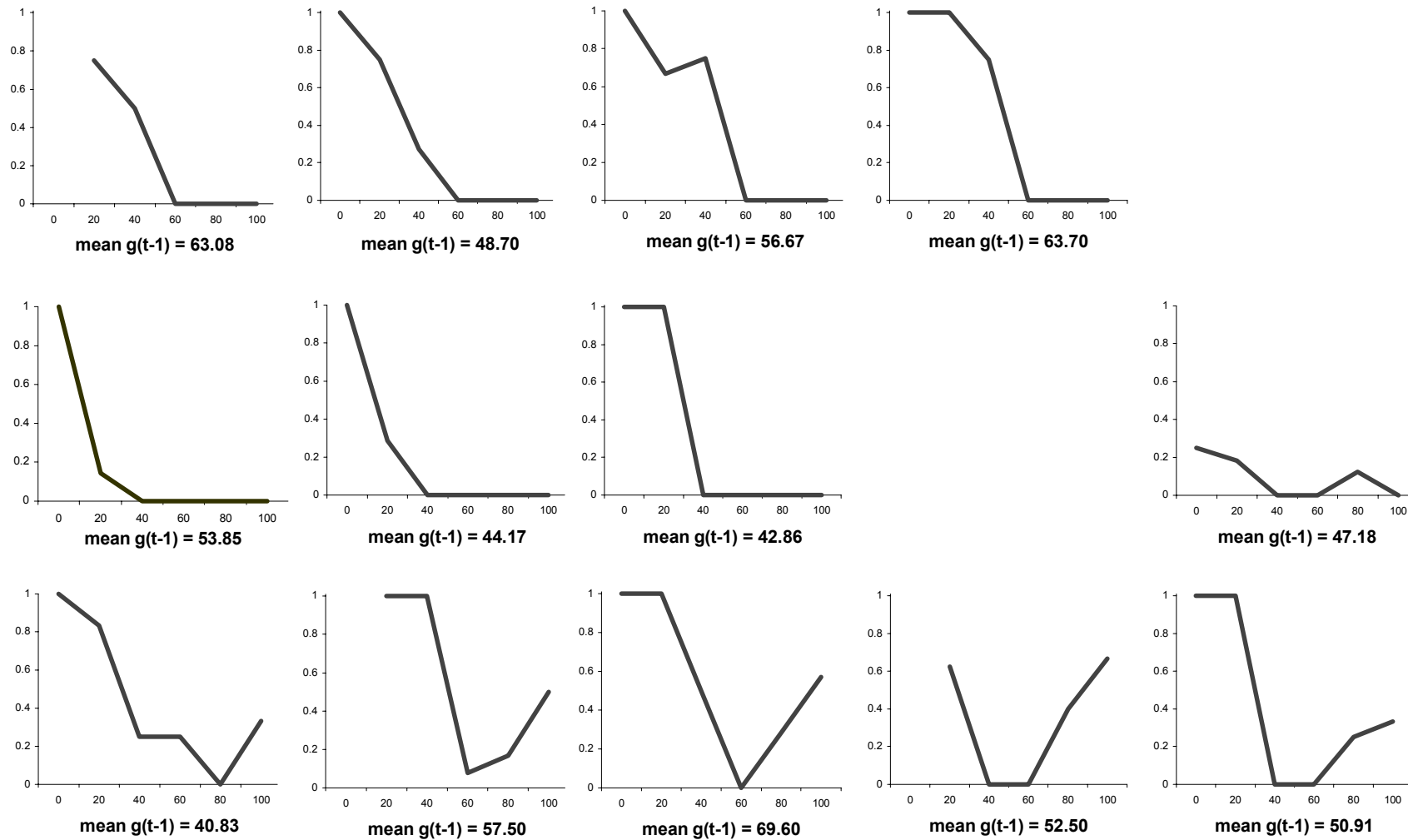


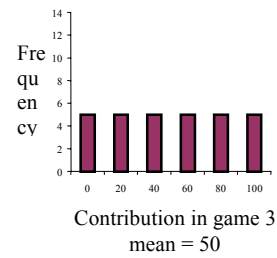
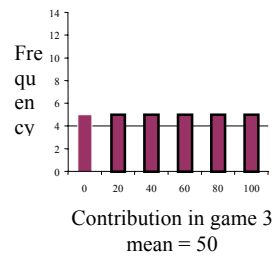
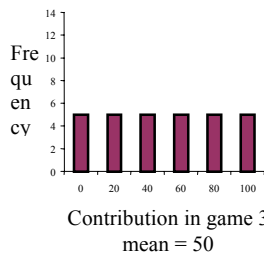
Figure 5: Simulated effects of witnessed and experienced criticism in three communities with different incidences of criticism

(a)

(b)

(c)

Contributions in game 3

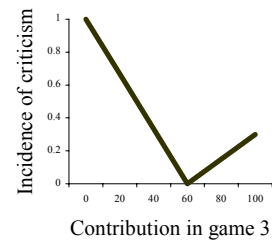
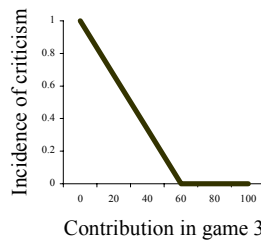
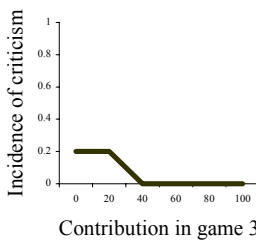


Incidence of criticism in discussion round after game 3

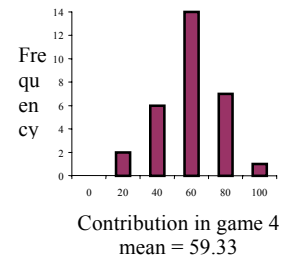
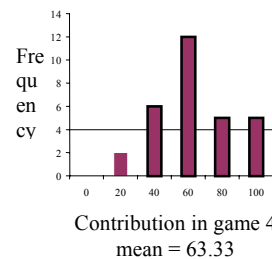
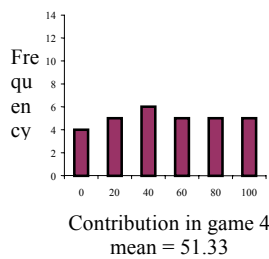
some criticism of low contributors

much criticism of low contributors

plus some criticism of high contributors



Contributions in game 4 taking account of impact of witnessing criticism only



Contributions in game 4 also taking account of the shemalessness of criticised nominees

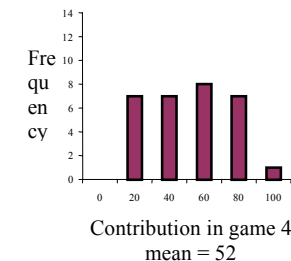
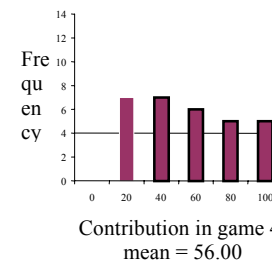
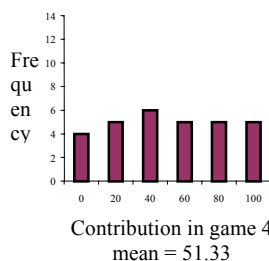


Table 1: Contributions by game and treatment in experiment 1

	Nominees who played anon-anon	Nominees who played anon-anon-public	All nominees	Spread of village means	Nominees in non-resettled village•
Mean g_{i1}	44.60	54.73***	50.65	35.00 – 63.00	57.00
Mean g_{i2}	53.00	55.81	54.68	38.00 – 68.00	59.00
Mean g_{i3}	51.00	60.81**	56.85	31.00 – 69.00	69.00
Observations	100	148	248		
Mean g_i	49.53	57.12***	54.06	31.00 – 69.00	61.67
Observations	300	444	744		20

** difference in means between treatments significant at 5% level, *** difference in means between treatments significant at 1% level, • mean contributions in non-resettled community do not significantly vary from mean contributions in resettled communities facing same treatment.

Table 2: Comparison of anonymous and public treatment, controlling for subject fixed effects(dependent variable = contribution (g_{it}))

	Full sample	$g_{it-1} = 0, 20$	$g_{it-1} = 40, 60$	$g_{it-1} = 80, 100$	Full sample
	(1)	(2)	(3)	(4)	(5)
	FE	FE	FE	FE	GMM
<i>constant</i>	26.324 (32.641)	112.179 (45.230)**	0.891 (37.778)	15.008 (64.977)	- -
g_{it-1}	-	-	-	-	-0.138 (0.373)
\bar{g}_{jt-1}	2.806 (2.985)	-8.232 (4.488)*	6.844 (3.488)*	3.128 (6.232)	3.318 (3.163)
\bar{g}_{jt-1}^2	-0.091 (0.098)	0.284 (0.160)*	-0.250 (0.118)**	-0.037 (0.209)	-0.110 (0.102)
\bar{g}_{jt-1}^3	0.001 (0.001)	-0.004 (0.002)*	0.004 (0.002)**	-9.47e ⁻⁵ (0.003)	0.001 (0.001)
\bar{g}_{jt-1}^4	-5.80e ⁻⁶ (6.37e ⁻⁶)	2.00e ⁻⁵ (1.18e ⁻⁵)*	-1.69e ⁻⁵ (7.98e ⁻⁶)*	2.64e ⁻⁶ (1.40e ⁻⁵)	-7.09e ⁻⁶ (6.54e ⁻⁶)
<i>t3 anon.</i>	-1.108 (3.220)	3.992 (4.795)	-7.077 (6.306)	-22.910 (5.028)***	- -
<i>t3 public</i>	5.184 (2.487)**	12.135 (4.079)***	3.347 (4.955)	-2.121 (3.847)	5.271 (2.355)**
FEs sig level	0.1%	0.1%	0.1%	0.1%	-
Observations	496	142	195	159	248
Subjects	248	101	149	114	248
R ² •	0.024	0.296	0.197	0.404	-

Standard errors in parentheses, * significant at 10% level, ** significant at 5% level, *** significant at 1% level, • within R² reported for fixed effects regressions.

Table 3: Criticism and praise in the discussion round

	Full sample	$g_{it-1} = 0, 20$	$g_{it-1} = 40, 60$	$g_{it-1} = 80, 100$
Criticised:	29.4%	67.7%	16.0%	14.0%
- for a low contribution	24.9%	67.7%	16.0%	0.0%
- for a high contribution	4.5%	0.0%	0.0%	14.0%
Praised	65.3%	29.2%	75.0%	83.3%
No comment	5.4%	3.1%	9.0%	2.6%
Observations	354	96	144	114

Table 4: Probit analysis of incidence of criticism (Dependent variable= *criticised*)

	Marginal effects	s.e.
g_{it-1}	-0.033	(0.047)***
g_{it-1}^2	$2.27e^{-4}$	$(3.95e^{-5})$ ***
village1	-0.261	(0.039)***
village2	-0.242	(0.036)***
village3	-0.210	(0.048)**
village4	-0.106	(0.121)
village5	-0.051	(0.160)
village6	-0.044	(0.107)
village7	0.023	(0.123)
village8	0.087	(0.149)
village9	0.110	(0.164)
village10	0.128	(0.162)
village11	0.307	(0.162)**
village12	0.342	(0.154)**
Pseudo R ²	0.4231	
Observations	354	
Joint sig. of village dummies	0.1%	

Standard errors in parentheses, * significant at 10% level, ** significant at 5% level, *** significant at 1% level

Table 5: Contributions by game in experiment 2

	Contributions	Spread of village means	Significance of variation between villages	Contributions by resettled nominees	Contributions by non-resettled nominees
Mean g_{i1}	44.35	35.83 – 64.62	10%	45.79	40.42
Mean g_{i2}	51.36	42.67 – 66.15	-	52.20	49.05
Mean g_{i3}	52.93	42.86 – 63.70	5%	53.44	51.58
Mean g_{i4}	51.02	35.00 – 75.38	0.1%	52.97	45.68*
Observations	354			258	96
Mean g_i	49.92		0.01%	51.10	46.68**
Observations	1416			1036	380

* difference in means for resettled and non-resettled nominees significant at 10% level, ** difference in means for resettled and non-resettled nominees significant at 5% level.

Table 6: Comparison of anonymous and post-discussion treatment
(dependent variable = contribution (g_{it}))

	Panel of all subjects involved in the second experiment		
	(1) FE	(2) FE	(3) GMM
<i>constant</i>	67.408 (13.163)***	68.299 (12.917)***	4.554 (1.750)***
g_{it-1}			-0.184 (0.052)***
\bar{g}_{jt-1}	-1.426 (0.874)	-1.425 (0.858)*	-1.468 (0.840)*
\bar{g}_{jt-1}^2	0.034 (0.018)*	0.033 (0.018)*	0.033 (0.017)**
\bar{g}_{jt-1}^3	-2.40e ⁻⁴ (1.19e ⁻⁴)**	-2.29e ⁻⁴ (1.17e ⁻⁴)*	-2.24e ⁻⁴ (1.07e ⁻⁴)**
<i>public</i>	3.775 (1.957)*	3.634 (1.921)*	
<i>public-ns</i>	-10.160 (3.289)***	-9.388 (3.241)***	-8.139 (2.663)***
<i>post-crit</i>	-3.253 (2.022)	-3.605 (2.134)*	-12.663 (2.987)***
<i>criticised-low</i>	11.960 (3.399)***	-11.959 (6.351)*	-12.438 (7.375)*
<i>criticised-high</i>	-15.059 (7.068)**	4.778 (8.802)	3.363 (7.199)
<i>prop-crit-low</i>		30.607 (7.489)***	38.852 (8.731)***
<i>prop-crit-high</i>		-46.379 (14.371)***	-53.744 (11.808)***
FEs sig level	0.1%	0.1%	
Observations	1062	1062	708
Subjects	354	354	354
R ² •	0.048	0.087	

Standard errors in parentheses, * significant at 10% level, ** significant at 5% level, *** significant at 1% level, • within R² reported for fixed effects regressions.