Maastricht University

UNITED NATIONS UNIVERSITY
UNU-MERIT

Working Paper Series

#2018-011

Sanctioning Regimes and Chief Quality: Evidence from Liberia

Gonne Beekman, Eleonora Nillesen and Maarten Voors

Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)
email: info@merit.unu.edu | website: http://www.merit.unu.edu

Maastricht Graduate School of Governance (MGSoG)
email: info-governance@maastrichtuniversity.nl | website: http://www.maastrichtuniversity.nl/governance

Boschstraat 24, 6211 AX Maastricht, The Netherlands
Tel: (31) (43) 388 44 00

UNU-MERIT Working Papers
ISSN 1871-9872

Maastricht Economic and social Research Institute on Innovation and Technology
UNU-MERIT

Maastricht Graduate School of Governance
MGSoG

UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at UNU-MERIT and MGSoG to stimulate discussion on the issues raised.

Sanctioning Regimes and Chief Quality: Evidence from Liberia

Gonne Beekman[*]
Eleonora Nillesen[†]
Maarten Voors[‡]

This version: February 2018

Abstract

We investigate how different sanctioning regimes and the quality of local leaders affects public goods provision in Liberian villages. We conduct a public goods experiment where leaders act as third-party punishers under one of two exogenously imposed sanctioning regimes. Under the first "flat fee" regime leaders receive a flat fee as compensation but do not receive any monetary gains from punishment. Under the second, "incentivised" regime leaders receive the punishment (tokens taken from the game participants) as compensation. We use villagers' perceived measures of corruption of their leader as our preferred measure for leadership quality. To empirically distinguish between sanctioning itself and the identity of the person sanctioning we have a treatment variation where a random villager acts as the third party punisher. We find that real village leaders elicit higher contributions than random villagers or groups without sanctioning. We also report that the effectiveness of sanctioning is attenuated by chiefs that are perceived to be of low-quality, especially when the sanctioner has no material incentive to punish. This suggests that low-quality chiefs are less likely to exert effort for public goods if they do not also privately benefit from it. Finally we find that people's preferred regime choice seems to depend on their real-life experiences in the village rather than their individual characteristics. Current development programmes heavily rely on community self-management and local institutions. Our paper supports the idea that a programme's success is likely to depend on whether villagers deem their leader to be credible norm enforcers.

Key words: Corruption, Public goods, Monitoring, Sanctioning, Field Experiment

JEL CODES: C9, K42, O17, O12

[*] Corresponding author. Development Economics Group, Wageningen University & Research. gonne.beekman@webmail.wur
[†] UNU-MERIT
[‡] Development Economics Group, Wageningen University & Research.

1 Introduction

The well-known problem of collective action, in which individual interests compete with the maximisation of social welfare, can be partially resolved by the establishment of sanctioning regimes. Such regimes create incentives, including pecuniary ones and social pressure, which steer individual behaviour towards welfare maximising outcomes. But what happens if these institutions are controlled or captured by elites who rig the system to favour their personal interests over the public good? This may be particularly salient in developing countries' settings where informal power relations are typically strong and leaders are not necessarily held accountable for their actions.

Indeed, in various social settings, there is an important symbiotic relationship between leaders and their underlings. For example, in many rural African communities chiefs play a central role, influencing daily life of villagers. Current popular participatory development projects heavily rely on the community's ability to self-organise and self-regulate investments in public goods. However, we know little about the ways in which chiefs motivate members of their community to invest in such public goods, or how they may sanction members of their community who deviate from proscribed social norms. Nor do we know to what extent leaders' actions correlate with perceived legitimacy of the leader.

To investigate these questions, we have designed a lab-in-the-field public goods game (PGG) with third-party sanctioning. We relate investment decisions and sanctioning behaviour in the PGG to sanctioning institutions and leadership quality. Subjects from rural villages in Liberia play multiple rounds of PGGs with and without third-party sanctioning by the village chief or by a randomly selected villager.[1] We want

[1] See Berg et al (1995) for a classic description of the public goods game.

to understand to what extent intrinsic quality of the chief may be compromised by institutional settings that create an opportunity for self-enrichment. We therefore randomly assign half of our village sample to either a "flat fee" or "incentivised" regime. Under the "flat fee" third party sanctioners receive 300 LD (about US$ 4.5), independent of their sanctioning behaviour. Under the incentivised regime the sanctioner is allowed to keep the tokens he decides to subtract from the players' accounts as punishment for the player's behaviour in that round. These tokens then become the sanctioner's private earnings. To measure chief quality we use average villagers' perception about corruption of their local leader. In addition we employ a measure similar to the one we used in Beekman et al., (2014): the chief's willingness to appropriate resources meant for the village in a structured community activity. Yet contrary to the our earlier study, our measure of appropriation has not been measured accurately; in about 30% of the villages we find the volume of rice to have increased rather than decreased after three days of storage. Although this theoretically could imply that the village chief added rice to the common sack, we consider this unlikely. Rather, we believe that in contrast to iron tools and wrapped packages of seeds as used in the Beekman et al. study, the stored rice may have been exposed to humid weather conditions (not unlikely in July in Liberia) that exogenously caused the volume to increase. We therefore use our "perception" measure (that is positively and significantly correlated with "rice missing") as our preferred measure of leader quality.[2]

We test five hypotheses. First, we expect the threat of third party sanctioning positively affects contributions in a PGG. We also expect the village chief to have greater leverage over his villagers than a randomly chosen leader. Thus village leaders are on average able to elicit larger contributions from their members than a randomly chosen villager. Second, we expect that leaders operating under an incentivised regime have an

[2] We also estimate models with the "rice missing" variable as key independent. Effects are both smaller and weaker.

additional private incentive to punish, in contrast to those operating under a "flat fee" where sanctioning must be due to non-monetary reasons including preferences for efficiency or equity. We hence expect chiefs to punish more under the incentivised regime, yet we do not necessarily expect this punishment to be effective (i.e. contributions do not necessarily increase given the chief's private incentive for punishment). Third, we test whether the effectiveness of sanctioning (i.e. eliciting higher contributions) varies with the quality of the leader. Fourth, we test if the effectiveness of sanctioning depends on the type of institutional regime and whether this interacts with the quality of local leadership.

All players play five rounds of each game A (PGG with village leader as sanctioner); B (PGG with random villager as sanctioner) and C (no sanctioning regime). The order of the games is randomised within villages. After 15 rounds have been played each player is invited to select the game of their choice (A, B, or C) and play their preferred game for another 10 rounds. In hypothesis five we then test whether contributions are different in groups where the regime (sanctioning or not, and by whom) was exogenously imposed than in groups that self-selected into such a regime. Finally, we look at punishment patterns to investigate whether chiefs exert more punishment to maximise equity or efficiency.

We find strong evidence that sanctioning by the village chief matters in the public goods game, more so than sanctioning by a random villager. We also find that leaders operating under a "flat fee" elicit higher contributions than those under an incentivised or no sanctioning regime. The identity of the leader however does not seem to matter here. When we interact the leaders' identity with our measures of perceived leadership we find the effectiveness of sanctioning is strongly mitigated if the chief is perceived to be corrupt, or of if the sanctioner comes from a village where its members perceive their local leader to be corrupt. The similarity in results between the village leader and a

random villager may suggest leadership quality in a village extends to prevailing norms. Results are less considerably weaker for our experimental measure and confined to the chief as sanctioner. Perhaps somewhat surprisingly we also find that higher levels of perceived corruption are associated with higher contributions. Although speculative, we interpret this as a possible signal of strong leaders that are able to enforce cooperation through their patronage network. These effects seem to be driven by leaders operating under the "flat fee" regime. This may suggest that low-quality leaders do not care enough about non-monetary incentives to punish or pose a credible threat that they will punish norm violators. When examining the determinants of "regime choice" for the bonus rounds we do not find strong individual predictors, but people are less likely to choose to play under a sanctioning regime if they perceive their chief to be corrupt, but more likely to do so if they live in villages with an arguably strong chief, and if their village has been attacked by warfare in the past. We believe these results are consistent with the idea that people bring their "real-life experience" to the lab. These experiences may have made them realise it is beneficial to have strong norm enforcers as leaders. Patterns of punishment suggest that sanctioners care about equity more than efficiency. Absolute negative deviations from the mean are punished significantly harder, but there is no effect on those with an absolute negative deviation from the maximum contribution. The results seem particularly driven by sanctioners playing under the flat rate regime.

The paper is organised as follows. In section 2 we provide an overview of the literature on corruption, leadership and public good provision under various monitoring and sanctioning regimes. In section 3, we provide information on our experimental setting in rural Liberia, and introduce our experimental design and measure of leadership quality. In section 4 we provide a simple model and derive predictions, which we test in our results section 5. Section 6 concludes.

2 Related literature

This study speaks to three literatures. First, our results contribute to the growing (experimental) literature on the impact of sanctioning (or rewarding) institutions on group outcomes. A common finding in this literature is that an effective sanctioning system is an important determinant of the success of overcoming social dilemma situations (e.g. Ostrom, et al., 1992; Fehr and Gächter, 2000; 2002; Cinyabuguma et al., 2005). These studies all exogenously impose positive or negative sanctioning regimes and are hence unable to inform us whether people would, if given a choice, prefer to operate under such institutions.

More recent studies address this particular issue and allow for an endogenous choice of institutional arrangements. Botelho et al. (2005) for example extend the design by Fehr and Gächter (2000, 2002) by having participants play a public goods game with and without a sanctioning norm in place. After subjects have played ten rounds of both versions, experimenters increase the stakes of the game for one final round: they multiply the value of a token in the previous rounds by ten, and ask participants to vote for their preferred version of the game. They find that people have a strong preference for a no-sanctioning regime.

Gürerk et al. (2006) allow participants to self-select into playing a public goods game under a sanction-free regime or one with positive and negative sanctioning. Individuals are allowed to "migrate" (at no cost) after every round. Slightly more than one third of the sample initially chooses the sanctioning institution and pay-offs are significantly lower than in the sanction-free environment, consistent with Botelho et al. (2005). Yet, the Gürerk et al. design permits them to analyse whether this result holds up in the long run. It does not. Free riding rapidly leads to a collapse in cooperation over time, with almost all participants migrating to the sanctioning institution, increasing their individual contributions and sanctioning norm violators. Interestingly, participants who

switch do not necessarily maximise their own individual pay-off because they invest some of their money in sanctioning norm violators (the second-order free-riding problem). Indeed, the majority starts sanctioning immediately after having switched. The authors view this as support for the idea that people tend to conform to common behaviour.

Related, Sutter et al. (2010) use a public goods set-up with the same three types of institutions as in Gürerk et al. (2006). Yet, they experimentally vary the way in which the institution is determined: exogenously (i.e. predetermined by the experimenter) or endogenously by having a costly vote among its group members. They additionally vary the efficiency of the sanctioning or reward across treatments to gauge additional impacts on cooperation. Cooperation is significantly higher among groups in the endogenous treatment compared to the exogenous treatment, even when sanctions or rewards are less efficient in the former than the latter. Possibly, by choosing a specific type of institution group members want to signal they adhere to a particular social norm.

Kosfeld et al. (2009) allows for an endogenous institutional choice. In their study, sanctioning institutions either form or they do not, depending on voluntary individual choices to participate in an organisation (institution) or not. They find that when organisations are formed (in about 50 percent of the cases), these institutions—although costly—have a positive impact on contributions in a public goods game. They are better able to maximise total welfare than in control groups where the possibility to form an organisation was precluded.

A small set of economic experiments examines the effect of third party sanctioning on contributions in lab games. Third party punishment provides perhaps the best measure of social norms as self-interest cannot confound normative preferences here. Fehr and Fischbacher (2004b) allow an unaffected third party to punish players in a dictator game and a prisoners' dilemma game, at a cost to the third party and excluding

any self-interest as a motivation for sanctioning. They find that approximately two-thirds of third party sanctioners enforced distributive norms by imposing punishments. Third party punishers were much more likely to punish a defector in a prisoners' dilemma game if the other player had cooperated, as compared to when the other player defected, showing that third party punishers have conditional perceptions of what is fair (Fehr and Fischbacher, 2004a). Third party punishment has a significant positive effect on the amount shared and the amount returned in the investment game (Charness et al., 2007). In this study the third party incurs a cost of punishment and gets no benefit other than the enforcement of a preferred norm for trust and trustworthiness in the investment game. Third party punishers are more prone to punish members of their own social group if they defect from social norms, than they are to punish members of a different social group for the same defection (Mizuho et al, 2004). In all of these studies, the third party punisher incurs a cost to punish, and the punishment is hence entirely altruistic. Also, experimental studies examining the impact of sanctioning or rewarding institutions are all conducted among college students in Western high-income countries with no evidence from the field. Our study is, to the best of our knowledge, the first to examine endogenous institutional choice in a field setting in a developing country. This has implications for the PGG design as we explain below.

In the context of a village in Liberia, the chief acts as an enforcer of social norms for the appropriate amount of public good provision, and has the ability to punish at no cost to himself. In fact, the chief can take resources from the village and use them for his own enrichment. We therefore add to the literature on third party sanctioning by introducing in some treatments a sanctioning regime, in which the third party not only incurs zero costs of punishment but even gains by doing so, and thus has a private material incentive

to punish. This set-up is, in our view, a more accurate representation of village life in Liberia, and possibly other rural African settings.

Second, our findings advance the literature on the role of leadership in economic development. Studies on leadership often investigate the management of firms and its relation to economic performance (e.g. Bertrand and Schoar, 2003; Malmendier and Tate 2005a, 2005b; Bloom and Van Rheenen, 2006). A notable example is Jones and Olken (2005) who examine the impact of national leaders on economic growth, exploiting death from natural causes or accidents as a source of exogenous variation in leadership transition. They find that leadership matters for growth in societies characterised by autocratic regimes and particularly within autocratic regimes that put fewer constraints on the power of their leader. A few theoretical studies demonstrate the importance of leadership in collective outcomes (Olson, 1965) or the evolution of social norms (Acemoglu and Jackson, 2015) and some of these predictions are now tested empirically. Acemoglu et al. (2014) for example collected historical data on ruling families in Sierra Leone from local elders and report that chieftaincies with fewer ruling families (i.e., less competition for power) are associated with worse human capital outcomes (e.g., literacy, child health, and the share of non-agricultural employment) but higher levels of *bridging* and *bonding* social capital, and more respect for authority within the chieftaincy. The authors interpret these two latter findings as evidence of powerful chiefs building social capital to monitor and control their citizens, while citizens strategically invest in patron-client relationships with their chiefs. Yet, they are unable to cleanly test this interpretation.

Chattopadhyay and Duflo (2004) make use of a national experiment in India in which a random subset of village council chief positions are reserved for women to study the consequences of the reservation system on the type of public goods provided.

Female leaders appear to invest more in public goods associated with women's preferences. The study however does not directly link leadership characteristics other than gender to the quality of public goods provided.

Khwaja (2009) reports a significant positive influence of the presence and quality of community leadership on maintenance of public infrastructure in Northern Pakistan. The authors measure leadership quality as the mean evaluation of five randomly selected household representatives within the community that would classify their leader as "good" or "bad".

Lastly, Kosfeld et al. (2015) examine sanctioning behaviour in a public goods game for a sample of real-world leaders of forestry management groups in Ethiopia and link the experimental play to observed outcomes in forestry management. Based on other-regarding preferences of leaders they identify a variety of leader "types" ranging from no sanctioners, leaders who sanction out of equality and (or) efficiency concerns and so-called 'anti-social' sanctioners who sanction co-operators or who's sanctioning behaviour is unrelated to contributions. Compared to non-sanctioners, sanctioners who are concerned about equality and efficiency have better forest management outcomes while the reverse applies to groups with leaders that apply anti-social sanctioning. Survey data reveal that the latter type of sanctioners are also more likely to be identified as "bad" leaders by their own group members, and were more likely to have been dismissed later, due to poor performance.

Finally, we contribute to the experimental economic literature on corruption.[3] Abbink and Serra (2012) provide a comprehensive overview of the effectiveness of various anti-corruption policies in the lab, including changing monetary incentives to take part in corrupt activities, and appeals to intrinsic motivations to abstain from

[3] There is an extensive non-experimental economics literature focusing on determinants (e.g. La Porta et al., 1999; Treisman, 2000; Serra, 2006) and (or) consequences (e.g. Mauro, 1995; Fisman, 2001; Fisman and Svensson, 2007; Olken and Barron, 2009) of corruption.

corrupt behaviour. An increasing number of scholars use field experiments to (i) assess the impact of various monitoring instruments to reduce it (e.g. Bjorkman and Svensson 2009; Olken 2007; Ferraz and Finan 2011, 2012; Callen and Long 2015) or (ii) use novel approaches to measure corruption and its consequences (e.g. Olken 2007; Bertrand et al., 2007; Beekman et al. 2014; Callen and Long 2015). Our study falls within the latter category, extending the design by Beekman et al. (2014), to which we turn next.

3 Experimental set-up and data

Experimental and survey data were collected in May-August 2012. We collected data among 2,560 individuals in 128 villages, spread out across three counties. We used a two-stage clustered stratified sample with village-level stratification based on participation in a rural livelihoods programme, implemented by an international NGO. Within each selected village we randomly subsampled 19 household representatives plus the village chief who were invited to participate in the household survey and the public goods experiments.

3.1 Measuring capture

We assume that the proclivity of a chief to capture resources for his own personal benefit when those resources are meant for others in the community is a behavioural measure of corruptibility that should be negatively correlated with leadership quality. We measure the proclivity of a chief to capture such resources using survey responses (aggregated at the village level) to the question whether a village member perceives their chief to be corrupt.[4] In addition we use a framed field experiment, akin to

[4] The specific question to the individual respondent read as follows: Generally speaking, what is your opinion of the likelihood of corruption when involving _____ (the president of Liberia/the district commissioner/the village chief/the village elders). Choice categories included none, low, medium, high, very high). To measure perceived corruption of the village chief we used aggregated responses to the subquestion that pertains to the village chief.

Beekman et al. (2014).[5] The field experiment was implemented as follows. Three days before conducting the household surveys and experiments the research team would visit each village to announce the upcoming research activities in the village and to hand a loosely sealed bag with 30 kilograms of rice to the chief. The chief was informed that the team would return after three days to conduct the experiments and household surveys, and that the rice would then be distributed to the villagers as a token of appreciation for taking part in the research activities. The chief was asked to safely store the rice in his private hut for this period of three days. On the day of the research activities, after performing the public goods experiments, the bag of rice was publicly weighed and divided into 20 small bags. After participating in the household survey, each participant (including the chief) received her share of the rice. Our measure of capture is the difference between the amount of rice distributed and the amount of rice left after three days for each chief in our sample.[6]

On average, 0.8 kilograms of rice disappeared during the three-day period of storage at the chief's hut, which corresponds to 3 percent of the total amount distributed (see summary statistics in Table 1).[7] We define leaders to be 'corrupt' (involved in capturing community inputs, measured in binary terms) if their bag was missing more than the median amount of rice relative to our sample of villages. Following this definition, 29 percent of the chiefs are labelled as corrupt.

[5] The key difference between a field experiment and a lab-in-the-field experiment, is that in the former, subjects are unaware they participate in an experiment. The other major difference is that field experiments have participants engage in an activity that they would ostensibly do in their day-to-day lives even without researchers showing up, whereas the activities performed as part of a lab-in-the-field experiment are less likely to be something the participant would do on her own. We were careful to frame the experiment such that the chief and the villagers would not get the impression that we were measuring the chief's behaviour. No one was informed about the amount of rice the bag was supposed to contain. The exact amount of rice was measured in private by the research team, just before handing the bag to the chief, in order to control for eventual spills during transportation or packaging.

[6] The exact amount of rice was measured in private by the research team, just before handing the bag to the chief, in order to control for eventual spills during transportation or packaging.

[7] Note that in 47 villages more rice was reported after storage than on the moment of distribution (0.4 kilogram on average). This is either due to measurement error, or, albeit less likely, chiefs added inputs.

<<Insert Table 1 about here>>

3.2 The public goods experiment

We invited subjects to play a public goods game (PGG) that included multiple rounds under different sanctioning regimes. In the first stage, we explained the rules of the PGG and randomly allocated players to one of three game versions.[8] Version A is a PGG with third party sanctioning by the village chief. After each round, and after the group learns how many tokens each anonymous player contributed and how many tokens each anonymous player currently has, the chief could punish any player in the group by taking any number of tokens from that player, up to the total number of tokens the player has at the time. Sanctioning is costless. This mimics the real-world power of the chief in his village, as he is able to claim resources from village members at no cost to himself. Each player is subsequently privately informed whether, and by how much, he was sanctioned. We limit reputational effects by announcing sanctioning decisions in private, only to the players who are sanctioned. Version B only differs from version A in having a different sanctioner: Version B has a randomly selected villager as sanctioner, in order to allow us to disentangle the possible confounding effects of sanctioning and the identity of the sanctioner. Version C is a regular PGG without third-party sanctioning. The game is played with the same group of six people for five rounds. In each round, players decide to allocate 15 tokens either to their private account or to the group account. Any token allocated to the private account is worth 10 LD to the player, and any token allocated to the group account is multiplied by 2 and equally distributed among all six players in the group (meaning each contributed token is worth 3.33 LD to each player in the group). After each round the experimenter publicly announces to the group the contributions of each group member as well as total earnings during that round.

[8] See Appendix X for detailed experimental set-up and instructions.

After the fifth round of their first version of the PGG, players are re-randomised into new groups and play five rounds of the second and third version of the PGG each. Each player hence plays all three versions of the game, but in different orders.

In addition, villages are randomly assigned to one of two incentive schemes for sanctioners in Version A and B, stratified on population size. In villages with incentive scheme 1, sanctioners receive a flat rate of 300 LD (about 4.5 USD) no matter how many times or by which amount they sanction their group members. Any tokens taken from the players' accounts during the sanctioning process return to the experimenters. In villages with incentive scheme 2, the sanctioner receives no additional salary beyond the 50 LD show-up fee that each player received. Yet, the sanctioner gets to keep any tokens he takes from players in his group. Hence, in these villages sanctioners have a private incentive to sanction. In both incentive schemes we explained to the group that the sanctioner was allowed to punish by taking tokens. It was framed as a sanctioning mechanism, but we did not restrict how the sanctioner used this power. It is possible that players viewed it not as punishment but as a justifiable tax that the sanctioner had a right to take. These two incentive schemes allow us to identify motives underlying the behaviour of sanctioners. Do they use their power in the village to increase public good provision because they care about the welfare of people in their village (hence they punish our of fairness or efficiency concerns) or do they rather use their power to sanction people if they can capture a private benefit from sanctioning (self-interested punishment)?

The second stage starts after the first ten rounds of play have been completed. All twenty participants in the village gather and the experimenter summarises the differences between the three game versions and informs everyone about the average amount players earned in each version. Participants are then invited to play another set of ten rounds but can now freely choose under which institutional environment they

would want to play. Only the sanctioner is unable to change groups. Groups now consist of minimal three and maximal eighteen players. If fewer than three players choose a specific game version, that version is disbanded and participants choose a second choice regime to join. Any token allocated to the group account is now worth 20 LD. The total value of all tokens contributed to the public good is divided equally among the players that joined the group.

Table 2 reports summary statistics of contributions, sanctioning behaviour and earnings from the PGG, by game version and sanctioning style. On average, players contribute about 8.7 tokens to the group account, which equals nearly 60 percent of their total endowment. Sanctioning size is 1.4 tokens on average and total earnings are 22.8 tokens. Overall, contributions and earnings are higher, and sanctions are lower, under incentive scheme 1 (flat rate) than under incentive scheme 2 (incentivised sanctioning). Contributions are highest under chief sanctioning and lowest without sanctioning. However, because some of the earnings are taken away in the game with third-party sanctioning, final earnings are highest in the game without sanctioning.

<<Insert

Table 2 *about here*>>

3.3 Household survey

After completing the game we conducted household surveys. We asked game participants four questions related to the game that preceded the actual household survey. The household questionnaire included standard questions about demographics and socio-economic variables, and a module on respondents' perception of corruption. To make the question on perceived corruption of the local chief (our key variable of interest) less sensitive and hence less susceptible to biased responses, we start out by asking questions about corruption at the national, county-and district level, before asking respondents to answer this question for their local village chief. Table 1 reports summary statistics of household characteristics. The average household head is 43 years of age. Nearly 80 percent of the household heads are male, 52 percent are born in the village, and 52 percent are literate. On average, households are comprised of 5 members. Households own 5.5 different types of assets, and in the past year they farmed 4.7 acres of land and hired 3.4 labourers. 22 percent of households sold food items in the capital city Monrovia in the past 6 months. In the previous month, total expenditures were 8.64 and -2.10 in logs in the previous 6 months. Finally, 38 percent of households report being attacked during the civil war.

4 Model, predictions and results

4.1 Institutional environment and contributions

We test if third party sanctioning affects contributions in the public goods game. We expect the threat of sanctioning to elicit greater contributions, as has been documented elsewhere. Let G_i be the public goods contribution made by player *i*,

regardless of the game version. G_i^{PC} is the contribution under sanctioning by the chief and G_i^N is the contribution in the game version that has no third-party sanctioning. We test if the average contributions across game versions are significantly different from each other.[9] Models are estimated at the group-round level with household and community controls and district fixed effects. All models control for the of tokens sanctioned in the previous round. Standard errors are clustered at the village level.

Hypothesis 1: Contributions are highest in the PGG with third-party sanctioning by the chief, and lowest in the PGG without third-party sanctioning.

$$G_i^{PC} > G_i^{PV} > G_i^C$$

To test Hypothesis 1, we estimate the following model:

$$G_{ij} = \alpha + \beta_1 P_{ij}^C + \beta_2 P_{ij}^V + \beta_3 r_{ij} + \beta_4 X_{ij} + \beta_5 C_j + \beta_6 \delta_d + \varepsilon_{ij} \quad (1)$$

where a group's contribution G_i in round r in village j is explained by dummies for the game version (P^C is sanctioning by the chief, P^V is sanctioning by the random villager, and the game without third-party sanctioning is the excluded base-variable). We control for the round of the game r_{ij} which ranges between 1 and 5 in the first five rounds, and between 6 and 15 in the ten bonus rounds. Our vector of household controls (X_{ij}, aggregated at the group-level) includes characteristics of the household head (age, sex, born in village, literacy), household size, assets owned, size of farmland, hired labour, sale of food items in the capital, expenditures, savings, and incidence of war attacks. Our vector of community controls (C_j) includes village size, connection to the grid, accessibility via an all-weather road, the presence of a weekly market in the village, NGO

[9] We use both average contributions over the five rounds of each version of the game, and the contribution in the last of five rounds after learning has occurred.

activity, and the share of villagers relying on agriculture as prime source of income. We also include district fixed effects (δ_d).

Results are presented in Table 3, columns (1) and (2). Having the chief as sanctioner invites higher contributions relative to having a random villager as sanctioner or no sanctioning regime at all. Having the chief as sanctioner (random villager) increases contributions by about 10 (8) percent compared to having no sanctioning in place. The null hypothesis of equality of coefficients for the chief and random villager is rejected ($p<0.10$) in two of the four models. The bonus rounds show less variation in contributions across the different game versions.

<<Insert Table 3 about here>>

Next, we turn to the sanctioning incentives. We expect that if sanctioners can keep the money they take from others, this will have an effect on the probability and size of sanctioning. With a flat rate payment, sanctioners have no private incentive to sanction, but they can sanction for other reasons including preferences for efficiency or fairness.

Hypothesis 2: If incentivised sanctioning leads to less well-targeted sanctioning, then sanctioning is less effective and contributions will drop. If incentivised sanctioning leads to higher overall sanctioning (level effect), but not to less targeted sanctioning (distribution effect), then contributions will be unaffected.

To test Hypothesis 2, we estimate the following two models:

$$G_{ij} = \alpha + \beta_1 P_{ij}^C + \beta_2 P_{ij}^V + \beta_3 F_j + \beta_4 r_{ij} + \beta_5 X_{ij} + \beta_6 C_j + \beta_7 \delta_d + \varepsilon_{ij} \qquad (2a)$$

and

$$G_{ij} = \alpha + \beta_1 P_{ij}^C + \beta_2 P_{ij}^V + \beta_3 F_j + \beta_4 P_{ij}^C \cdot F_j + \beta_5 P_{ij}^V \cdot F_j + \beta_6 r_{ij} + \beta_7 X_{ij} +$$

$$\beta_8 C_j + \beta_9 \delta_d + \varepsilon_{ij} \quad (2b)$$

where all variables are defined as in (1) and F_j is a dummy taking the value 1 if chiefs operate under a non-incentivised ("flat fee") sanctioning regime.

Results in Table 1 column (3) show that sanctioners under a flat rate elicit higher contributions compared to those in the incentivised regime, but again only in the first five rounds (columns 1 and 2). The insignificance of the interaction term indicates that the identity of the sanctioner does not matter when operating under the flat fee (column 4).

4.2 Leadership quality, institutional environment and contributions

Our second set of hypotheses deals with the effects of leadership quality on behaviour in the PGG. Some chiefs may be more effective or credible sanctioners than others. We would like to explore the heterogeneous effects of leadership quality on public good provision. Specifically, we equate more capture and higher levels of perceived corruption with lower leadership quality.

Hypothesis 3: We expect a low-quality chief to be less able to elicit high contributions than credible, "good" leaders.

To test Hypothesis 3, we estimate the following model:

$$G_{ij} = \alpha + \beta_1 P_{ij}^C + \beta_2 P_{ij}^V + \beta_3 Q_j + \beta_4 P_{ij}^C \cdot Q_j + \beta_5 P_{ij}^V \cdot Q_j + \beta_6 r_{ij} + \beta_7 X_{ij} +$$

$$\beta_8 C_j + \beta_9 \delta_d + \varepsilon_{ij} \quad (3)$$

where all variables are as defined in (1) and Q_j is our measure for chief quality – either survey measure (perceived corruptibility) or the experimental (amount of rice missing).

Table 4 presents the results. Column (1) indicates that the positive impact of having the chief as sanctioner is mitigated if this is considered a low-quality chief.

Contributions are considerably higher under sanctioning regimes but the effect almost halves in villages where chiefs are considered corrupt. This effect is even stronger for the villager-as-sanctioner. Our experimental measure (column 2) is much weaker and only significant for the chief as sanctioner. This may be due to the fact that the rice stored was unambiguously connected to the chief, whereas perceived corruptibility may have had a broader applicability.

In Table 5 we estimate the same models but split the sample according to the different sanctioning regimes (flat or incentivised) and only present results for our preferred survey measure of corruption. The results in table 4 seem to be largely driven by groups operating under the flat fee regime. A low-quality leader affects contributions only if this leader receives a flat fee (effects are smaller and not significant in the incentivised subsample). This may suggest that low-quality leadership especially turns out "bad" if these leaders operate under an institutional regime that appeals to demonstrating intrinsic leadership qualities like preferences for fairness, equity and (or) efficiency.

4.3 Game choice for the bonus rounds

What then happens when the game version is endogenously chosen? Recall that our experiment includes ten final rounds for which participants self-selected into their regime of choice. We included these rounds to estimate public good outcomes based on revealed preference, and to learn more about how institutions may evolve over time in such villages. We expect leader quality to play a role in their decision about which version to play—in villages with a high-quality leader who caused players to contribute more, players would be more likely to select into version A with the chief as sanctioner.

In villages with a low-quality leader, players may avoid the chief's group either because they feel his sanctioning is arbitrary and serves as a way for him to enrich himself at their expense, or because he does not command the kind of respect that leads to higher contributions from opponents. We run two separate regression models, one with a dummy for choosing a sanctioning regime (Table 6, column 1) and a second one with a dummy for chief sanctioning (Table 6, column 2) to assess whether the identity of the sanctioner matters. The second model is estimated using the data only from participants choosing a sanctioning regime. People are less likely to choose a sanctioning regime if they perceive the sanctioner to be corrupt or come from a corrupt environment. This is consistent with our hypothesis that when given a choice, people prefer not to operate under bad quality institutions. Also, individuals living in villages with large plantations, (which suggest that in these villages the chief has more power), are more likely to choose to play under sanctioning regimes just as those living in villages that have been attacked by the war in the past. One interpretation may be that people bring their real-life experience to the lab. People living in villages with powerful chiefs may recognise the benefits of having a strong norm enforcer. This may then explain why game choice is not determined by individual traits but village history. Contributions are on average however lower in the final ten (bonus) rounds than in the previous rounds where people played according to the regime exogenously imposed on them (table 7). Yet this effect is due to the low contributions that are made by people that choose to play under a non-sanctioning regime. Also contributions are lower in the game versions with sanctioning but this effect is offset when players self-select into a sanctioning game.

Our final sets of results are related to the actual sanctioning in the game. We here replicate the Fehr and Gachter (2000) model, where our dependent variable is the number of tokens sanctioned. The relevant independent variables are mean investment

levels, absolute negative deviations from the mean and max (sanctioning for equity or efficiency reasons) and our measures of corruption. We here find a negative correlation between mean investment levels and sanctioning, consistent across most rounds and sanctioning type (flat versus incentivised). Higher negative deviations from the mean are systematically related to higher sanctioning, suggesting the sanctioners in our sample did have a preference for equity. Yet results seem to be largely driven by sanctioners receiving a flat rate. This is consistent with the idea that bad quality leaders exert less effort for the greater good if they do not also personally gain from it.

5 Discussion

Our results show support for many but not all hypotheses tested in this experiment. For Hypothesis one, we compare contributions across the two main versions of the public goods game (no sanctioning and sanctioning by the chief) and find that in games with the chief as sanctioner contributions are higher compared to having a random villager as sanctioner or no sanctioning at all, consistent with our hypothesis that real-world leaders are better able to enforce norms than leaders randomly chosen in a game.

For our second hypothesis, we looked at the different sanctioning regimes and find that sanctioners receiving a flat rate are able to elicit higher contributions than those operating under an incentivised regime. This provides support for the idea that private material incentives may (partially) crowd out leaders' non-material motives for public good provision.

We next hypothesise that chief quality, using survey-based and experimental measures affects contributions. We indeed find that low-quality leadership mitigates the positive impact of a sanctioning chief. This effect is most pronounced for the perceived rather than the experimental measure of corruption. These measures are positively and

significantly correlated, suggesting that both give us some sense about corruption and arguably quality of local leadership. There may be multiple reasons why we find a stronger result using the survey response not the experimental measure of corruption. First, as noted, in some 30 percent of the villages we find more rice after three days, rather than less suggesting either measurement error due to e.g. exposure to weather conditions like humidity or (less likely) chiefs adding rice (for reasons unknown to us). Second, the stronger results on perceptions may indicate that what matters is people's beliefs about their leader rather than his actual behaviour.

Finally, we use data from the last ten rounds when players were able to choose their own regime to learn more about preferences for sanctioning under different institutional regimes. We find that game choice is affected by perceived levels of corruption: people are less likely to choose a sanctioning regime if they perceive their leader to be corrupt. By contrast, villagers with a powerful leader and those that have experienced a war attack in their community are more likely to choose a sanctioning regime. This may suggest people feel they need strong norm enforcers to protect them from harm. None of the other variables turn out significant, suggesting that real-life common experiences (war) or the day-to-day institutional environment (low-quality leadership) rather than innate characteristics, determine people's choice of institutions in the game.

6 Conclusion

As we think about the efficiency of social institutions at helping individuals reach socially optimal outcomes, we are interested in understanding the obstacles that may prevent groups from generating potential surplus. While much empirical evidence has been found to support the idea that institutional regimes can improve efficiency, in real-

world situations where social groups are tightly connected and regime leadership has strong private incentives, we hypothesise that sanctioning regimes operate less effectively. We used outcomes from a public goods experiment, survey responses and a framed field experiment to detect differences in regime performance depending on the identity and quality of the leader.

Our results support most of our hypotheses and point us towards future areas of research. We find strong evidence that sanctioning by the chief enhances contributions and that this result somewhat extends to having a random villager as sanctioner.

We also find that low-quality leaders attenuate the positive impact of sanctioning institutions. We use two measures of chief quality: the experimental measure of "missing rice" leads to considerable less strong and less robust results than using perceived measures of corruption from survey responses. In contrast to the Beekman et al. paper where we use a similar experimental set-up to measure corruption, we believe the use of rice in the current paper instead of iron tools and packaged seeds may have caused measurement error that led to finding more rather than less rice in about 30 percent of the villages. The stronger result on the perception measure may also indicate that what matters are people's beliefs about their leader rather than his actual behaviour.

This may have important implications for development policy. The newest generation of development programmes heavily rely on community self-management and building local institutions. Yet, our results show that leaders ability to enforce social norms through sanctioning institutions strongly co-varies with people's perception of their leader. Policymakers should moreover think hard about the set-up of these institutions. Providing material incentives to local leaders may on the one hand motivate them to exert punishment and hence enforce norms; yet may on the other hand (partially) crowd out leaders' innate preferences for equity or efficiency, and thus compromise the provision of public goods.

References

Abbink, K., Serra, D., 2012. Anticorruption Policies: Lessons from the Lab. In: Serra, D., Wantchekon, L. (Eds.), New Advances in Experimental Research on Corruption. Research in Experimental Economics, vol. 15. Emerald, Bingley, pp.77–115.

Acemoglu, D., Reed, T., Robinson, J.A., 2014. Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone. *Journal of Political Econonmics* 122(2), pp.319-368.

Acemoglu, D., M. Jackson, 2015 History, Expectations, and Leadership in the Evolution of Social Norms. *Review of Economic Studies* 82(2), pp.423-456.

Beekman, G., Bulte, E., Nillesen, E., 2014. Corruption, investments and contributions to public goods: Experimental evidence from rural Liberia. *Journal of Public Economics* 115(1), pp.37-47.

Bertrand, M., A. Schoar, 2003. Managing with Style: The Effect of Managers on Firm Policies. *The Quarterly Journal of Economics* 118 (4), pp.1169-1208.

Bertrand, M., Djankov, S., Hanna, R., Mullainathan, S., 2007. Obtaining a Driver's License in India: An Experimental Approach to Studying Corruption. *Quarterly Journal of Economics* 122(4), 1639–1676.

Bjorkman, M., Svensson, J., 2009. Power to the People: Evidence from a Randomized Field Experiment of Community-Based Monitoring in Uganda. *Quarterly Journal of Economics* 124(2), 735–769.

Bloom, N., Van Reenen, J., 2006. *Measuring and explaining management practices across firms and countries*. No. w12216. National Bureau of Economic Research.

Botelho, A., Harrison, G. W., Pinto, L., & Rutström, E. E., 2005. *Social norms and social choice*. Working Paper no. 30, University of Miho and NIMA.

Callen, M., Long, J. D., 2015. Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan. *American Economic Review* 105(1), pp.354-381.

Chattopadhyay, R., Duflo, E., 2004. Women as Policy Makers: Evidence from a India-Wide Randomized Policy Experiment. *Econometrica* 72 (5) 1409-1443.

Cinyabuguma, M., Page, T., Putterman, L., 2005. Cooperation under the threat of expulsion in a public goods experiment. *Journal of Public Economics*, *89*(8), 1421-1435.

Fehr, E., Gächter, S., 2000. Fairness and retaliation: The economics of reciprocity. *The journal of economic perspectives*, 159-181.

Fehr, E., Gächter, S., 2002. Altruistic sanctioning in humans. *Nature*, *415*(6868), 137-140.

Ferraz, C., Finan, F., 2011. Electoral Accountability and Corruption: Evidence from the Audits of Local Governments. *American Economic Review* 101, 1274–1311.

Ferraz, C., Finan, F., Moreira, D.B., 2012. Corrupting learning: Evidence from missing federal education funds in Brazil. *Journal of Public Economics*. 96, 712–726.

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B., 2006. The competitive advantage of sanctioning institutions. *Science*, *312*(5770), 108-111.

Jones, B.F., Olken, B.A., 2005. Do Leaders Matter? National Leadership and Growth Since World War II. *Quarterly Journal of Economics* 120, 835–864.

Kosfeld, M., Okada, A., Riedl, A., 2009. Institution formation in public goods games. *The American Economic Review*, 1335-1355.

Khwaja, A.I., 2009. Can Good Projects Succeed in Bad Communities? *Journal of Public Economics,* 93, 899–916.

Kosfeld, M., Rustagi, D., 2015. Leader Sanctioning and Cooperation in Groups: Experimental Field Evidence from Commons Management in Ethiopia. *American Economic Review*, 105(2), pp.747-783.

Malmendier, U., & Tate, G., 2005a. CEO overconfidence and corporate investment. *The journal of finance*, *60*(6), 2661-2700.

Malmendier, U., & Tate, G., 2005b. Does overconfidence affect corporate investment? CEO overconfidence measures revisited. *European Financial Management*, *11*(5), 649-659.

Olken, B., 2007. Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy* 115, 200–249.

Olson, M., 1965. *The logic of collective action: Public goods and the theory of group* (p. 176). Cambridge: Harvard University Press.

Ostrom, E., Walker, J., Gardner, R., 1992. Covenants with and without a Sword: Self-governance Is Possible. *American Political Science Review*, *86*(02), 404-417.

Sutter, M., Haigner, S., Kocher, M.G., 2010. Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *The Review of Economic Studies*, *77*(4), 1540-1566.

Table 1: Summary statistics

Variable	Obs	Mean	SD	Min	Max
Panel A: Capture					
Perceived corruption	130	0.85	0.46	0.1	2.3
Amount of rice missing	128	0.81	1.72	-1.6	9.1
Corrupt (b)	128	0.29	0.45	0	1
Panel B: Household controls					
Age head in years	2561	42.85	14.47	16	102
Male head (b)	2553	0.79	0.41	0	1
Household size	2546	5.07	2.27	0	16
Head born in village (b)	2561	0.52	0.50	0	1
Literate head	2494	0.52	0.50	0	1
Assets owned	2512	5.48	2.55	1	17
Land size (acres)	2588	4.65	29.07	0	800
Sold food in Monrovia in the past 6 months (b)	2439	0.22	0.41	0	1
Number of hired laborers	2548	3.42	9.10	0	251
Log Total savings	2485	-2. 10	8.81	-9.21	13.82
Log Total expenditures	2485	8.64	2.21	-9.21	12.82
War attack (b)	2456	0.38	0.49	0	1

Note: (b) binary outcome, 1 = yes

Table 2: Summary statistics PGG

Variable	Version 1 Incentivised		Version 2 Flat Fee		Diff.	t-test
	Obs	Mean (SE)	Obs	Mean (SE)		p-value
Game A: Chief	882	9.08 (0.08)	930	8.80 (0.08)	0.29 (0.12)	0.01
Game B: Villager	858	8.80 (0.09)	843	8.70 (0.08)	0.10 (0.12)	0.41
Game C: None	857	8.78 (0.10)	894	8.33 (0.09)	0.45 (0.14)	0.00
Sanctioning chief	882	1.20 (0.04)	930	1.51 (0.04)	-0.31 (0.05)	0.00
Sanctioning villager	857	1.37 (0.04)	843	1.51 (0.04)	-0.14 (0.06)	0.02
Game A: Earnings	882	22.87 (0.09)	930	22.29 (0.09)	0.58 (0.13)	0.00
Game B: Earnings	858	22.43 (0.10)	843	22.35 (0.11)	0.08 (0.15)	0.58
Game C: Earnings	857	23.78 (0.10)	894	23.34 (0.10)	0.44 (0.14)	0.00

Table 3: PGG contributions and sanctioning regime

	Round 1-5		Round 1-5	
	(1)	(2)	(3)	(4)
Chief sanctioning	0.423***	0.403***	0.407***	0.556***
	(0.149)	(0.151)	(0.151)	(0.196)
Villager sanctioning	0.185	0.196	0.192	0.442*
	(0.164)	(0.166)	(0.166)	(0.243)
Flat rate			0.590*	0.867**
			(0.313)	(0.423)
Chief × Flat rate				-0.315
				(0.327)
Villager × Flat rate				-0.525
				(0.350)
Round	0.196***	0.192***	0.192***	0.192***
	(0.034)	(0.035)	(0.035)	(0.035)
Constant	7.812***	6.426***	6.104***	5.988***
	(0.217)	(1.850)	(1.814)	(1.808)
HH controls				
Comm. controls				
District FEs	No	Yes	No	Yes
No. Clusters	110	108	108	108
N	4995	4905	4905	4905
R^2	0.016	0.111	0.121	0.122
P-value test Chief = Vill.	0.08	0.13		

Note: Standard errors clustered on village level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: PGG contributions and leadership quality

	Round 1-5	Round 1-5
	(1)	(2)
Chief sanctioning	1.018***	0.748***
	(0.218)	(0.262)
Villager sanctioning	0.752***	0.353
	(0.237)	(0.300)
Perceived corruption	1.335***	
	(0.211)	
Rice missing		0.092
		(0.133)
Chief × Perceived corruption	-0.705***	
	(0.180)	
Villager × Perceived corruption	-0.641***	
	(0.214)	
Chief × Rice missing		-0.139*
		(0.073)
Villager × Rice missing		-0.053
		(0.101)
Round	0.192***	0.182***
	(0.035)	(0.034)
RA dummy	0.027	-0.113
	(0.357)	(0.397)
Constant	6.144***	5.744***
	(1.568)	(1.818)
HH controls		
Comm. controls		
District FEs	Yes	Yes
No. Clusters	108	106
N	4905	4830
R^2	0.151	0.114

Note: Standard errors clustered on village level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: PGG contributions, sanctioning regime & perceived corruption

	Round 1-5	
	Flat rate	Incentivised
	(1)	(2)
Chief sanctioning	0.996**	0.847***
	(0.417)	(0.281)
Villager sanctioning	0.758*	0.702**
	(0.413)	(0.313)
Perceived corruption	0.969***	1.309***
	(0.325)	(0.279)
Chief × Perceived corr.	-0.842**	-0.399+
	(0.325)	(0.243)
Villager × Perceived corr.	-0.914***	-0.348
	(0.322)	(0.295)
Round	0.255***	0.155***
	(0.054)	(0.050)
RA dummy	0.572	-0.764*
	(0.425)	(0.453)
Constant	7.727***	5.700***
	(2.788)	(2.001)
HH controls		
Comm. controls		
District FEs	Yes	Yes
No. Clusters	47	56
N	2140	2405
R^2	0.169	0.282

Note: Standard errors clustered on village level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Game choice

	Sanctioning group	Chief sanctioning
	(1)	(2)
Total amount sanctioned	0.001	0.005
	(0.004)	(0.007)
Perceived corruption	-0.086***	-0.005
	(0.030)	(0.049)
Age head	0.002	-0.002
	(0.002)	(0.003)
Male head	-0.047	-0.019
	(0.089)	(0.107)
HH size	0.021+	0.025
	(0.014)	(0.021)
Born in village	0.004	0.069
	(0.074)	(0.095)
Literate head	0.099	0.005
	(0.076)	(0.105)
Asset count	-0.013	0.013
	(0.015)	(0.020)
Land size	0.001	-0.000
	(0.001)	(0.002)
Sold food in capital	0.142+	0.048
	(0.092)	(0.123)
Hired labour	0.002	0.004
	(0.004)	(0.004)
Log savings	0.001	0.003
	(0.004)	(0.005)
Log expenditures	-0.010	0.001
	(0.019)	(0.024)
War attack	0.286***	0.084
	(0.070)	(0.095)
Comm. controls		
Flat rate	0.134+	0.084
	(0.084)	(0.117)
Community size	0.000	0.000
	(0.000)	(0.000)
Electricity	0.066	-0.519***
	(0.181)	(0.195)
Main road	0.030	0.213*
	(0.096)	(0.121)
Market	0.092	-0.227+
	(0.167)	(0.142)
NGO	-0.048	-0.110
	(0.094)	(0.113)
Share of agr. hhs	-0.001	0.002
	(0.002)	(0.002)
Plantation size	0.001**	-0.000
	(0.000)	(0.000)
Private land	0.194+	-0.252*
	(0.128)	(0.146)
RA dummy	0.044	0.293**
	(0.103)	(0.132)
Constant	-0.298	-0.642
	(0.320)	(0.464)
N	1632	984
Pseudo R^2	0.041	0.044

Table 7: Contributions exogenous versus endogenous game choice

	(1)	(2)	(3)
Bonus round	-0.541***	-0.671***	-0.929
	(0.169)	(0.173)	(0.268)
Round	0.121***	0.138***	0.118***
	(0.016)	(0.016)	(0.016)
Sanctioning (b)			-0.640***
			(0.309)
Bonus round × Sanctioning			0.667
			(0.326)
Constant	8.239***	7.229***	4.620***
	(0.163)	(1.695)	(1.470)
HH controls	No	Yes	Yes
Comm. controls			
District FEs			
N	7813	7693	7548
R^2	0.014	0.109	0.106

Note: Standard errors clustered on village level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Chief sanctioning (all rounds combined)

	All	Flat rate	Incentivised
	(1)	(2)	(3)
Bonus round	-0.060	0.003	-0.126
	(0.112)	(0.132)	(0.192)
Round	-0.009	-0.013	-0.007
	(0.010)	(0.010)	(0.018)
Mean investment	-0.183***	-0.166***	-0.189***
	(0.017)	(0.019)	(0.027)
Absolute negative deviation from the mean	0.405***	0.393***	0.436***
	(0.086)	(0.114)	(0.118)
Absolute negative deviation from the max	-0.089**	-0.111**	-0.094*
	(0.041)	(0.051)	(0.055)
Perceived corruption (b)	0.353*	0.258	0.366
	(0.207)	(0.329)	(0.265)
Corrupt × Abs. neg. deviation mean	-0.069	-0.146	-0.035
	(0.100)	(0.117)	(0.153)
Corrupt × Abs. neg. deviation max	0.067	0.127*	0.055
	(0.065)	(0.070)	(0.097)
Constant	2.986***	2.900**	2.206**
	(0.772)	(1.270)	(0.848)
HH controls	No	Yes	Yes
Comm. controls			
District FEs			
N	2675	1190	1315
R^2	0.288	0.368	0.265

Note: Standard errors in parentheses. $^{+}$ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (b) refers to a binary variable of perceived corruption that is 1 if people perceive any likelihood of corruption (from low-medium to high) and 0 otherwise.

APPENDIX

Table A1: Game choice for bonus rounds

Village code	Group A Chief sanctioning	Group B Villager sanctioning	Group C No sanctioning	Total
1	0	8	10	18
2	6	6	6	18
3	6	6	6	18
4	3	8	7	18
5	5	5	8	18
6	4	4	10	18
7	0	10	8	18
8	8	0	10	18
9	3	4	11	18
10	12	4	2	18
11	4	3	11	18
12	4	3	11	18
13	8	5	5	18
15	7	6	5	18
16	7	0	11	18
17	4	5	9	18
18	4	3	11	18
19	4	6	8	18
21	7	2	9	18
22	13	0	5	18
23	6	5	7	18
24	15	0	3	18
26	0	0	18	18
29	6	5	7	18
30	4	3	11	18
31	0	8	10	18
32	5	6	7	18
33	3	4	11	18
34	7	5	6	18
35	2	13	3	18
37	7	5	6	18
38	8	10	0	18
39	7	6	5	18
41	6	0	12	18
42	8	0	10	18
43	6	7	5	18
44	3	4	7	14
46	11	0	7	18
47	13	5	0	18
48	0	9	9	18
49	4	5	9	18
50	13	5	0	18
52	3	3	12	18
53	7	7	4	18
54	5	5	8	18
56	8	0	10	18
57	5	7	6	18
58	6	5	7	18
59	6	6	6	18
60	6	8	4	18
61	5	5	8	18
62	9	6	3	18
63	5	4	9	18
65	9	3	6	18
66	5	6	7	18
67	5	5	8	18

69	7	5	6	18
70	7	5	6	18
71	10	0	8	18
72	4	6	8	18
73	6	7	5	18
74	3	5	9	17
75	6	6	6	18
76	5	0	13	18
77	5	5	8	18
78	0	0	18	18
79	7	5	6	18
80	9	3	6	18
81	3	4	11	18
82	6	6	6	18
83	6	6	6	18
84	10	4	4	18
85	7	5	6	18
86	14	0	4	18
87	0	0	18	18
88	6	5	7	18
89	6	8	4	18
90	7	4	7	18
91	7	5	6	18
93	15	0	3	18
94	7	7	4	18
95	9	5	4	18
96	7	5	6	18
97	5	0	13	18
98	6	0	12	18
99	9	0	9	18
101	10	0	8	18
102	7	5	6	18
103	6	0	12	18
104	5	6	7	18
105	6	6	6	18
106	6	6	6	18
107	6	4	8	18
108	6	5	7	18
109	5	7	6	18
110	5	5	8	18
111	5	5	7	17
112	5	6	7	18
113	10	4	4	18
114	8	5	5	18
115	9	9	0	18
116	6	4	6	16
117	5	8	5	18
118	5	7	6	18
120	9	0	9	18
121	10	8	0	18
122	5	5	8	18
123	5	7	6	18
124	6	6	6	18
125	6	5	7	18
126	7	5	6	18
127	5	7	6	18
128	7	6	5	18
129	10	3	5	18
131	5	7	6	18
132	7	7	4	18
133	6	7	5	18

134	4	8	6	18
136	6	5	7	18
890	0	6	12	18
TOTAL	739	558	855	

The UNU-MERIT WORKING Paper Series

2018-01 *The serendipity theorem for an endogenous open economy growth model* by Thomas Ziesemer

2018-02 *Instability constraints and development traps: An empirical analysis of growth cycles and economic volatility in Latin America* by Danilo Sartorello Spinola

2018-03 *Natural, effective and BOP-constrained rates of growth: Adjustment mechanisms and closure equations* by Gabriel Porcile and Danilo Sartorello Spinola

2018-04 *The relevance of local structures for economic multiplier effects of social pensions in Uganda* by Maria Klara Kuss, Franziska Gassmann and Firminus Mugumya

2018-05 *Regulating the digital economy: Are we moving towards a 'win-win' or a 'lose-lose'?* by Padmashree Gehl Sampath

2018-06 *The economic impacts of a social pension on recipient households with unequal access to markets in Uganda* by Maria Klara Kuss, Patrick Llewellin and Franziska Gassmann

2018-07 *The effect of weather index insurance on social capital: Experimental evidence from Ethiopia* by Halefom Y. Nigus, Eleonora Nillesen and Pierre Mohnen

2018-08 *Evaluating intergenerational persistence of economic preferences: A large scale experiment with families in Bangladesh* by Shyamal Chowdhury, Matthias Sutter and Klaus F. Zimmermann

2018-09 *Agricultural extension and input subsidies to reduce food insecurity. Evidence from a field experiment in the Congo* by Koen Leuveld, Eleonora Nillesen, Janneke Pieters, Martha Ross, Maarten Voors and Elise Wang Sonne

2018-10 *Market integration and pro-social behaviour in rural Liberia* by Stephan Dietrich Gonne Beekman, and Eleonora Nillesen

2018-11 *Sanctioning Regimes and Chief Quality: Evidence from Liberia* by Gonne Beekman, Eleonora Nillesen and Maarten Voors