SUPPORTING ONLINE INFORMATION
Centralized-sanctioning and legitimate authority promote cooperation in humans

Delia Baldassarri,[1*] Guy Grossman[2]

[1]Department of Sociology, Princeton University; dbalda@princeton.edu
[2]Department of Political Science, Columbia University; gsg2102@columbia.edu

Contents

1 Supporting Analysis, Tables, and Figures

1.1 Analysis of contributions

In the analysis of contributions (see two-sided Mann-Whitney tests in the text) we argue that the difference between treatments and baseline (random *vs.* baseline; elected *vs.* baseline), as well as the difference between treatments (random *vs.* elected) are significant. Additional support for this finding comes from three statistical models in which we control for preliminary contributions and cluster the standard errors at the session level (Tab. S1). Contributions to the public account in public goods games are treated as proportions or fractions of the initial endowment (in our case a fraction of 1,000 USH). Using a proportion in linear regression models will generally yield out-of-range predictions for extreme values of the regressors. We follow Papke and Wooldridge's (1996) strategy for handling proportions in which zeros and ones may appear, as well as intermediate values. We apply a generalized linear model (GLM), using a logit transformation of the dependent variable and the binomial distribution. Results using Tobit models are very similar.

	Model A	Model B	Model C
Baseline-Random	0.22**		
	(0.08)		
Baseline-Elected		0.40***	
		(0.09)	
Random-Elected			0.17*
			(0.08)
Preliminary contribution	2.55***	2.24***	2.20***
	(0.15)	(0.15)	(0.15)
Constant	-1.38***	-1.26***	-1.02***
	(0.09)	(0.09)	(0.08)
Observations	990	970	926
Log Likelihood	-449.04	-447.36	-423.52
AIC	904.08	900.72	853.05

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tab. S1. Generalized linear model. DV: Contributions in round 3 to 6 (logit transformation). Clustered standard errors at the session level. Standard errors in parentheses. See text for further details.

1.2 Analysis of expectations and response to punishment

EXPECTATIONS	Coef.	s.e.
Elected monitor condition	39.230*	(18.64)
Monitor gender	1.493	(23.87)
Monitor born in village	-2.857	(19.19)
Monitor age	0.919	(0.72)
Monitor church attendance	6.878	(24.72)
Monitor education	-1.828	(3.57)
Monitor wealth	-1.557	(3.32)
Intercept	39.369*	(18.98)
Session level constant (σ_u)	23.849	(20.06)
Constant (σ_e)	237.737***	(5.95)
Observations	882	
Log likelihood	-6081.306	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tab. S2. Change in contribution due to the expectation induced by the monitor. Multilevel regression model. DV: Change in individual contribution. Grouping at the Session level. We model the change in individual contribution from preliminary round 2 to round 3 as a function of whether the sanction comes from an elected monitor, controlling for the monitor's sociodemographic characteristics. This table complements Fig. 2A in the text.

RESPONSE to PUNISHMENT	Coef.	s.e.
Elected monitor condition	-18.282	(12.45)
Sanctioned at $t-1$	57.293***	(16.04)
Sanctioned by Elected	60.840**	(21.47)
Preliminary contribution	-0.009	(0.02)
Number of Players in Session	-1.575	(4.37)
Number of players sanctioned at $t-1$	3.043	(3.67)
Maximum contribution sanctioned at $t-1$	-0.046	(0.03)
Minimum contribution sanctioned at $t-1$	-0.043	(0.03)
Monitor gender	-2.589	(14.70)
Monitor age	-0.193	(0.47)
Monitor education	-1.039	(2.23)
Monitor wealth	1.774	(2.03)
Monitor church attendance	0.676	(15.02)
Monitor born in village	2.561	(11.91)
Intercept	34.018	(50.19)
Observations	2448	
N Sessions	91	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tab. S3. Change in contribution as a reaction to sanctioning. Multilevel regression model. DV: Change in individual contribution from $t-1$ to t (i.e., from round 3 to 4, 4 to 5, and 5 to 6). Grouping at the session level. We model the change in player i's contribution from $t-1$ to t as a function of whether player i has been sanctioned at $t-1$, and whether the sanction comes from an elected monitor, controlling for the monitor's sociodemographic characteristics, monitor's sanctioning behavior at time $t-1$, and player i's preliminary contribution. This table complements Fig. 2B and 2C in the text.

1.3 Test of "leadership selection" effect

In the text we consider the possibility of a leadership selection effect. Fig. S1 shows the average socio-demographic characteristics of the monitors in the random and elected monitor conditions. Elected monitors are disproportionally male, locally born, more educated and richer. There are, however, no differences with respect to age and church attendance.

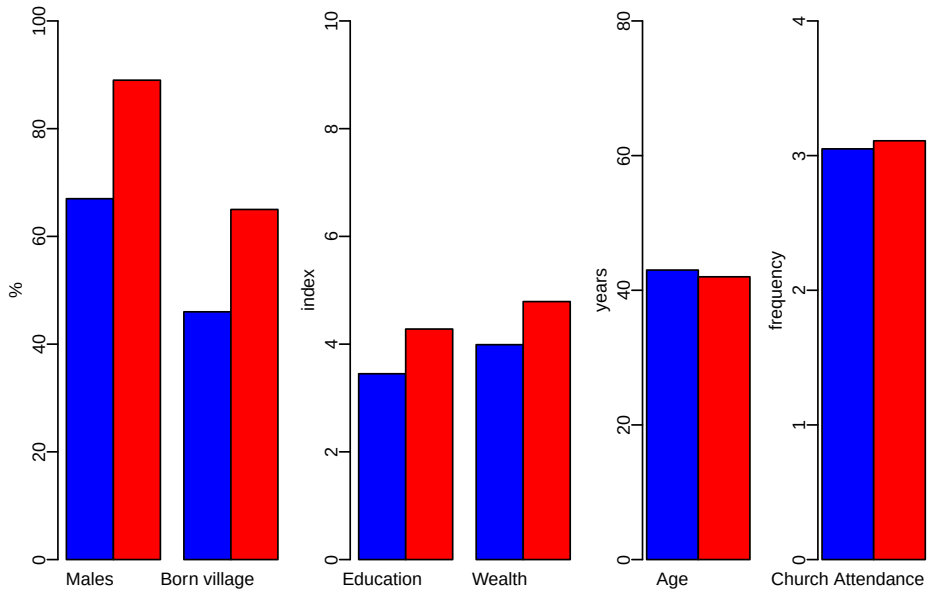


Fig. S1. Average sociodemographic characteristics of the monitors in the random (blue) and elected monitor (red) condition.

Tab. S4 shows that the monitor's profile does not affect contributions. We run two models to test the significance of the monitor's sociodemographic profile. The first is an OLS model and the second is a GLM model. In both cases, the dependent variable is players' mean contribution in rounds 3 to 6. In the GLM model, the mean contribution is calculated as a proportion of the maximum contribution (bounded by 0 and 1). In both models, the mean contribution is modeled as function of six characteristics of monitors (gender, local birth, age, church attendance, education and wealth). We report results from regressions run on data from the random treatment sample only. Since in the random treatment the monitors were randomly chosen, we can exclude the possibility of spurious correlations, and confidently assess the impact of monitors' sociodemographics on contributions. Moreover, since some of these sociodemographic characteristics are strongly correlated with each other (i.e., males are, on average, richer and more educated than women), we compute a joint significance test for the OLS model and consider two statistics. According to the R^2 statistics, the model explains a mere 5% of the variance in contributions. Similarly, the value for the F statistic is 1.34, which is much lower than the critical value needed for significance ($P = 0.2598$).

	OLS		GLM	
	Coef.	s.e.	Coef.	s.e.
Monitor gender	8.074	(42.06)	0.032	(0.17)
Monitor born in village	-27.473	(40.72)	-0.111	(0.16)
Monitor age	-0.983	(1.15)	-0.004	(0.00)
Monitor church attendance	88.567	(62.13)	0.359	(0.25)
Monitor education	-12.016	(6.75)	-0.048	(0.03)
Monitor wealth	-2.109	(5.13)	-0.009	(0.02)
Intercept	472.679***	(28.27)	-0.110	(0.11)
Observations	446		446	
R^2	0.049			
Joint Significance	1.34			
Degrees of Freedom	6.00		6.00	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tab. S4. Impact of the monitor's socio-demographic profile on contributions. OLS and GLM. DV: mean contribution in rounds 1 to 4. Players' contributions are modeled as a function of monitor's sociodemographic profile. Joint Significance Test. See text above for further details.

1.4 Analysis of monitors' behavior

In this section we provide a detailed analysis of monitors' behavior. We first provide a descriptive analysis of the data, and then present results from a matching analysis in which we take into account the fact that monitors faced different distributions of contributions.

1.4.1 Descriptive analysis

Elected and random monitors sanction, on average, the same number of players, from 2.5 players sanctioned in round 3 to 1.5 in round 6 (Tab. S5). Monitors differ in the range of contributions sanctioned. In the first round, both random and elected monitors sanctioned players at similar levels of contribution. In subsequent rounds, the maximum contribution sanctioned by random monitors gradually declined, while the maximum contribution sanctioned by elected monitors increased. By round 6, players in the random monitor condition who contributed more than 25% of their initial endowment were not punished, while players in the elected monitors condition were sanctioned for contributions up to 37% of their initial endowment In round 6 the difference in the maximum amount sanctioned is statistically significant ($P = 0.022$). We should nonetheless take into account the fact that monitors face different distributions of contributions.

Frequency of Punishment				
	Round 1	Round 2	Round 3	Round 4
Random Monitor	2.47	1.86	1.78	1.49
Elected Monitor	2.40	2.08	1.71	1.58
Minimum Punishment				
	Round 1	Round 2	Round 3	Round 4
Random Monitor	162.22	174.42	208.89	191.89
Elected Monitor	181.82	228.89	231.11	231.58
Maximum Punishment				
	Round 1	Round 2	Round 3	Round 4
Random Monitor	308.89	267.44	275.56	245.95
Elected Monitor	320.45	360.00	328.89	371.05

Tab. S5. Frequency, minimum, and maximum average punishment in the random and elected monitor condition. N=97 (49 random and 48 elected monitors)

In Fig. S2, we show, for each level of contribution, the proportion of contributions that were sanctioned. The data confirms that both types of monitors have similar sanctioning patterns: the likelihood of being sanctioned, under the two sanctioning treatments, drops sharply between 0 to about 400 USH, and then flattens for higher levels of contributions. Though there is some evidence suggesting that elected monitors tend to sanction a greater proportion of those contributions that fell in the 200-400 USH range, this difference is not significant. More so, this difference may be the result of the different distribution of contribution that both monitors face, from Round 6 onwards.

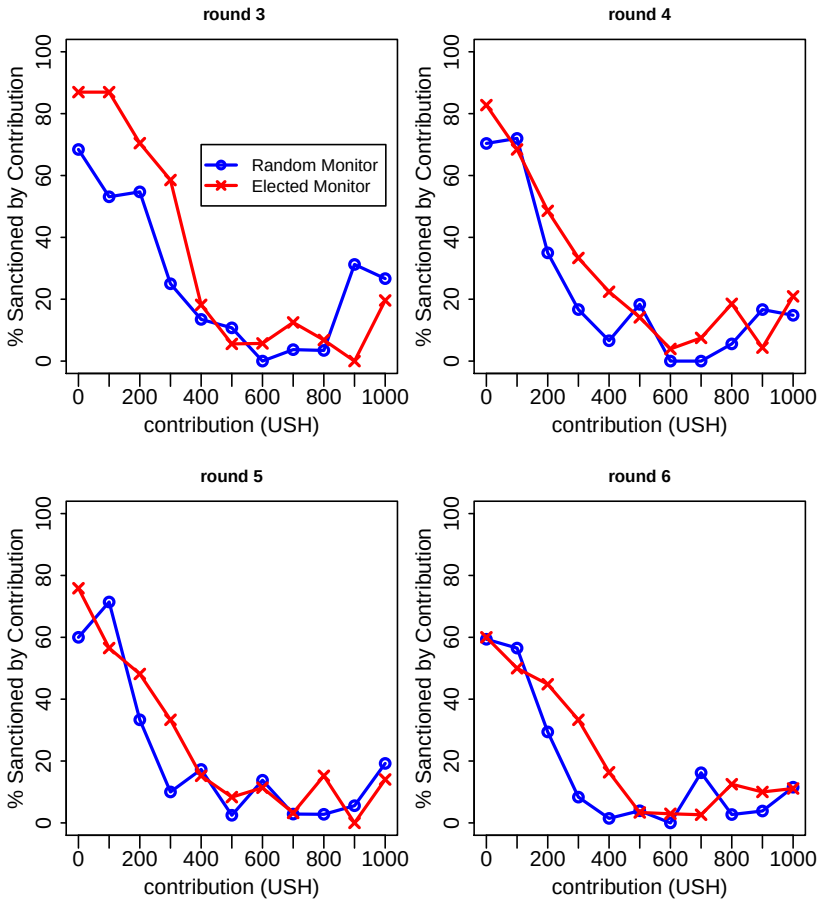


Fig. S2 Proportion of contributions sanctioned by level of contribution in the random monitor (blue) and elected monitor (red) condition.

Finally, Fig. S3 shows that there is no relationship between frequency of punishment and average group contribution.

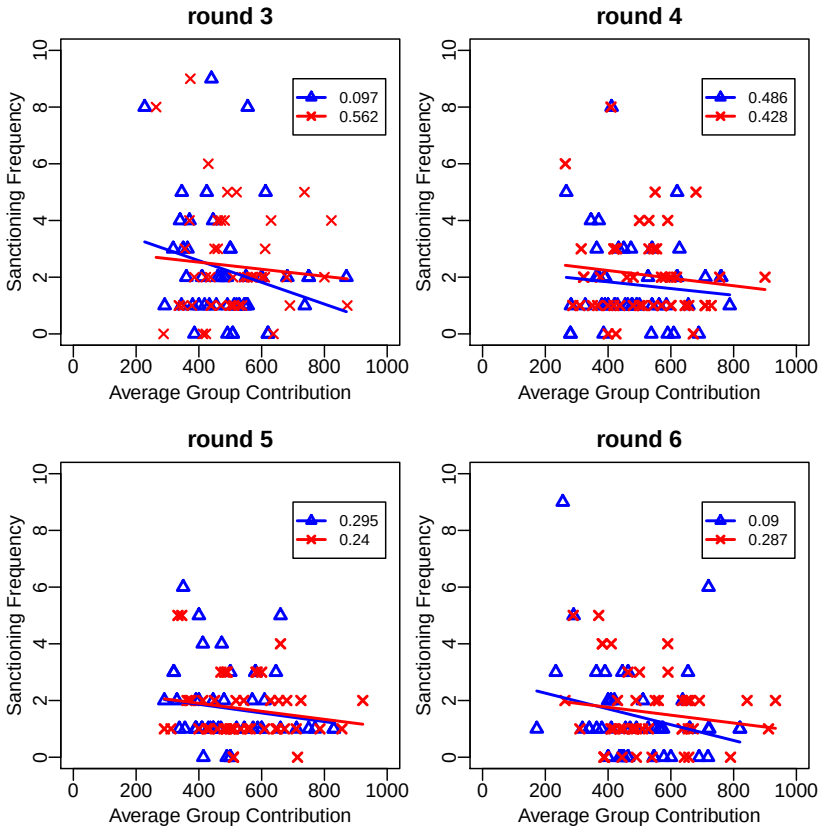


Fig. S3. Plot of the number of people sanctioned by average group contribution in the random (blue) and elected monitor (red) condition. Regression line and p-value reported in the plot area.

1.4.2 Matching analysis based on the distributions of contributions

To compare the behavior of the two types of monitors and circumvent the inferential problem stemming from the different distributions of contribution we matched each of the *distributions of contributions* to the public pot that an elected monitor faced with the closest distribution of contributions that a random monitor faced. Using the matched pairs, we can then test whether elected and random monitors facing similar distributions of contributions differ in their sanctioning strategies. Comparing the behavior of monitors within matched pairs, we find no evidence that elected monitors punished more frequently, nor that monitors used a higher threshold to signal an accepted level of cooperation.

First, we calculated for each elected monitor's distribution of contributions the distance to all possible random monitors' distributions, using the Kullback–Leibler (K–L) divergence measure. We then employed, for each round, a nearest neighbor with replacement matching algorithm. The Kullback–Leibler (K–L) divergence is a non-symmetric measure, commonly used to calculate the difference or distance between two probability distributions. For probability distributions P and Q of a discrete random variable, their K–L divergence is defined to be:

$$D_{KL} = (P||Q) = \sum P(i) log \frac{P(i)}{Q(i)} \tag{1}$$

The improvement in the divergence between distributions of contributions achieved by matching is substantial: on average, the matching algorithm reduces the K–L distance between the distributions that the different monitors faced, by a factor of about 4. Tab. S6 shows that in all sanctioning rounds, the mean K–L divergent measure of *the matched pairs* was similar.

		Matched Pairs				All T2–T1 Pairs	
	Mean	sd	Min	Max	N Pairs	Mean	N pairs
Round 3	0.091	0.050	0.027	0.228	48	0.328	2,352
Round 4	0.079	0.032	0.030	0.182	48	0.306	2,352
Round 5	0.088	0.041	0.025	0.178	48	0.336	2,352
Round 6	0.074	0.044	0.010	0.183	48	0.310	2,352

Tab. S6. K–L Divergence Measure for both Matched Pairs and for all possible T2–T1 pairs, by round.

We use two measures of behavior to test whether elected monitors enforce stronger norms of cooperation than random monitors: (i) the number of players being punished in Round t, and (ii) the maximum contribution being punished in Round t. The left panel in Fig. S4 graphs the number of players punished by elected monitors, as a function of the number of players punished by random monitors, for all matched pairs. Jittered dots that are above (below) the 45% line are matched pairs in which the elected monitor punished more (less) players than matched random monitor. Similarly, the right panel in Fig. S4 graphs the maximum contribution punished in Round t by elected monitors, as a function of the behavior of random monitors, for all matched pairs. We observe as many dots above the line than below the line, suggesting that there is

no difference between random and elected monitors with respect to the number of players and maximum contribution sanctioned. These results are summarized by Fig. 3 in the main text, in which we report differences in mean behavior of the matched pairs.

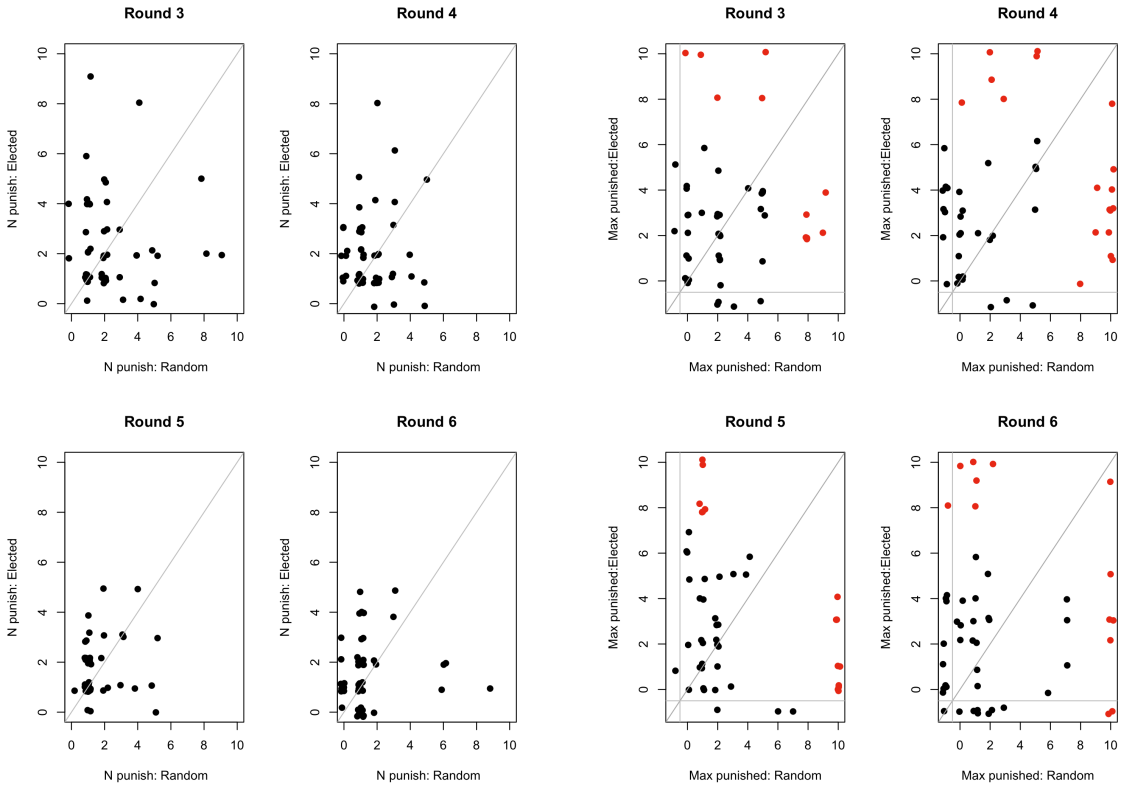


Fig. S4. Comparison of the sanctioning behavior of matched pairs of monitors. Plots show the A) Number of players punished in Round t by elected monitors as a function of the number of players punished by the random monitors, for all matched pairs; and B) Maximum contribution punished in Round t by elected monitors, as a function of the maximum contribution punished by the Random monitors for all matched pairs. Red dots signal that the maximum contribution punished in that round was equal or above 8 coins (MUs). Note that -1 on the X (Y) axis refer to sessions in which the random (elected) monitor has not punished any player (more accurately, any level of contribution.) Punishment decisions (x-axis) are reported in MUs (coins).

2 Broader Scope of the Research

Development scholars regard farmer organizations as a core component of poverty reduction strategies, but little is known about the social dynamics that make certain producer organizations more successful than others. In the past two decades, governmental withdrawal from direct assistance to small farmers has dramatically increased the importance of local-level, voluntary-based organizations in the provision of collective goods. The broader scope of our research is to study how producer organizations in rural Uganda solve classic problems of collective action, and determine the role of social and spatial networks, associational capital, and leadership accountability in affecting economic and social outcomes.

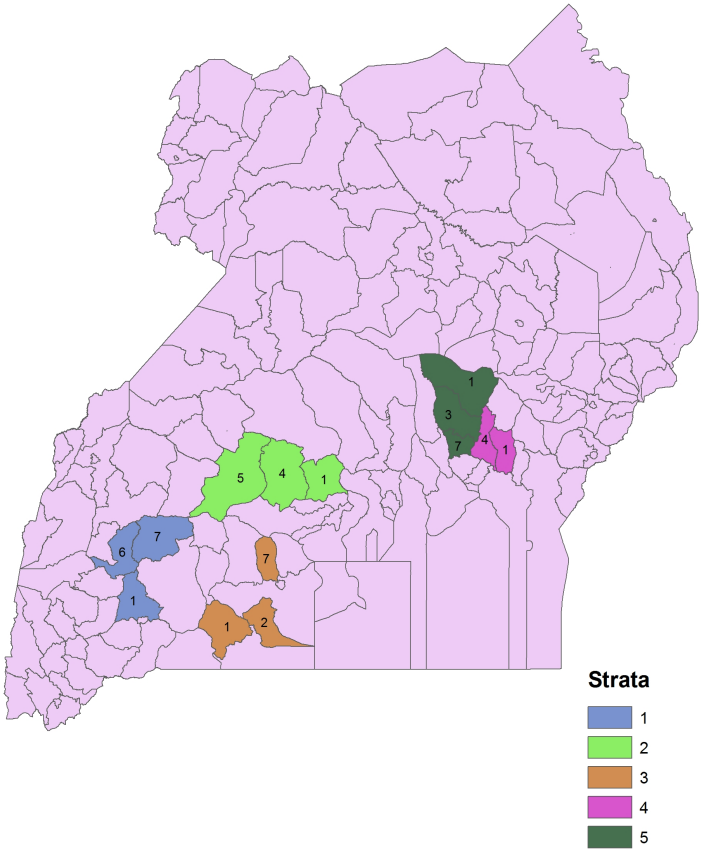
We focus on one of Uganda's largest rural developments project in recent years – the Agriculture Productivity Enhancement Project (APEP). Since 2005, APEP has helped organize over 60,000 small-holder farmers into more than 2,500 village-level farmer groups. Those village-level groups were further organized into larger farmer associations.[1] Serving about 200 members from ten neighboring village-level groups, farmer associations are better positioned to exploit economies of scale and to bargain for better prices based on larger volumes.[2] The stated goal of APEP is to expand rural economic opportunities by supporting subsistence farmers' transition into commercial farming and increasing small farmers' productivity and marketing capabilities. Its success varies across producer organizations, due to variation in local leaders' capacity to spread information, elicit relationships of trust, and facilitate the emergence of accountability practices.

Through a multilevel and multimethod research design, data have been collected at the farmer (i.e. group member), village group, and farmer association levels. From July to September 2009, our team of approximately 60 interviewers has gathered extensive information on a sample of 50 farmer associations via group interviews with local leaders, and has administered approximately 3,500 individual interviews with a sample of 'ordinary' group members as well as with group and association leaders. We have collected social network information for each of our subjects, and complete social network data for a subsample of 100 village-level producer organizations. In addition, both farmers and leaders have played an extensive series of behavioral games, among which was the public goods game presented in this paper.

[1] In APEP's jargon, village-level farmer groups who serve, on average, 20 members are referred to as Producer Organizations (POs), while the second-tier farmer associations are referred to as Depot Committees (DCs).

[2] In Uganda, districts are the most important administrative unit. Districts consist of 2-4 counties, each county has 3-6 sub-counties, each sub-county consists of 3-6 parishes, and each parish has about 5-15 villages. Between 200 and 800 households live in each village. DCs commonly serve one or two parishes, at most.

3 Sampling Strategy



We followed a stratified, multistage cluster design. Starting from 5 district-areas (strata), we sampled 50 parish-level farmer associations, then 6 village-level farmer groups and 36 farmers from each farmer association. We followed a five-step sampling strategy, summarized in the table below.

Step	Sampling Unit (SU)	Number of SUs	Sampling Method
1	Target Population	105 DCs	Coffee growers
2	District-area	5 strata	Stratified – proportional to # of DCs in strata
3	Farmer Associations	50	unequal probability without replacement
4	Producer Organizations (POs)	6 per DC	Clustered – simple random sample.
5	Group members	36 per DC	Clustered – probab proportional to POs size.

Step 1: Target Population. In order to reduce variability due to environmental factors, we limited our sample to farmer associations (DCs) that cultivate the same crop. Thus we included only coffee groups, as this was the most common crop. This decision reduced the universe of cases from 204 to 113 associations. An additional 8 DCs were excluded before sampling due to the following reasons: we excluded 2 associations from the Bugiri district because coffee is a

very unusual crop in that district. We further excluded 5 associations from the Busheni district because those groups were formed many years before APEP. Finally, we excluded from our sample a DC from the Kamwenge because it was the single DC in that district and surveying it would have been logistically complicated and prohibitively expensive. Our final universe comprises 105 coffee growing farmer associations, which were all founded after 2005.

Step 2: Strata. Based on power calculations performed on simulated data, we selected 50 associations, using a stratified random sample. Though our universe of farmer associations is spread over 9 districts, we grouped associations into 5 strata. Strata were defined by meaningful district-areas: neighboring districts that were covered by the same project field trainers and trading partners, and that share a dominant ethnicity and/or were historically part of the same district were grouped in the same strata (Fig. above).

Step 3: Farmer Associations. We draw independent probability samples of farmer associations from each stratum. We used unequal probability sampling without replacement to sample associations within strata (proportional to their size). The number of associations that were sampled from each stratum was proportional to the number of associations in each stratum. According to this scheme, sampled associations are representative at the stratum level.

Step 4: Village-level Farmer Groups. We want to be able to detect variation not only between associations but also within associations; i.e. in the success of village-level farmer groups that make-up each association. Thus, prior to sampling group members, we randomly sampled six farmer groups from each association. In the case that a farmer association had six or fewer village-level groups, we selected all of them.

Step 5: Random Sampling of Group Members. The same number of respondents (36) was sampled from each farmer association. The number of sampled members from each of the six *sampled* village-level groups was proportional to the size of the groups. This assured that the sample is self-weighted. Total sample size is thus 50 DCs $\times$ 6 groups $\times$ 6 members per group = 1,800. We succeeded in surveying 1,784 out of the 1,800 sampled group members (99% response rate), out of which 1,543 participated in the public goods games (86% of the subjects sampled).

3.1 Implementation

Data were collected between July 2009 and September 2009 by a group of 60 experienced local interviewers. Interviewers, who were hired directly by the PIs and their project manager, were divided into three "language" teams. The eastern team covered 16 farmer associations in the Iganga and Kamuli districts, where Basoga is the local language. The central team covered 20 DCs from the Mubende, Mityana, Masaka and Rakai districts, where locals speak Luganda. Finally, the western team covered 14 DCs from the Kiruhura, Mbarara and Ibanda districts, where Ranyankole is the lingua franca. Interviewers went through a lengthly training in class (4 days) and in the field setting (4 days), which included training on human subjects issues as well as survey techniques. Interviewers were supervised by team leaders: on average, there was a team leader for every 5 interviewers. This made it possible to constantly screen the work of the interviewers and greatly increase the quality of our data.

In each sampled association, data was collected in four rounds. First, an interviewer scheduled a meeting with the executives of the farmer association. In that meeting the interviewer

introduced the study to the leaders and asked for their cooperation. In that meeting the interviewer also administered the association-level questionnaire, as well as obtained a list of all DC council members. On the second day of interviews, the research team conducted interviews with group representatives to the DC council and with the chairmen of all village-level groups. All subjects were mobilized by the association's executives to a central location. In addition to the individual-level interviews, association leaders from each sampled village-group were asked to respond to a group-level questionnaire, and to provide a complete list of group members. Between the second and third days of interviews, the research team sampled 36 members from each sampled association (including 8 replacements). Immediately after the sampling procedure (see above), an interviewer travelled back to meet with the association leaders to hand them the list of sampled members and coordinate with them the next round of interviews. Once again, we relied on the associations leaders to mobilize the sampled members to a centralized location. On the third day of interviews, individual-level interviews were conducted with the sampled members and with representatives who were not present on the previous day. On that day we also conducted a set of behavioral games with both "ordinary" members and group representatives. Finally, the survey team traveled to each association for an additional day to interview sampled members or representatives who, for any reason, were not present in the previous days of interview.

3.2 Missing data

Great care was taken to reduce missingness. The research team administered association-level questionnaires in all 50 DCs and group-level questionnaires in 287 out of 289 sampled village groups. Out of a sample of 1,800 "ordinary" members (i.e., farmers with no leadership role in their group) the research team managed to conduct individual-level surveys with 1,784 farmers. Following standard practice for survey data missing at random (MAR), we used Patrick Royston's ICE multiple imputation package in Stata, which applies a chained equations approach. We imputed missing data for some asset variables and demographic characteristics, but chose not to impute data for farmers' agricultural practices.

3.3 Randomization procedure

Members' random allocation into the three game variants was conducted through the following procedure: each sampled member that arrived to the central location in the morning of day 3 (see above), first went through a registration and identity verification procedure. Following registration, each sampled member received a number in the order of arrival. The survey team simply took batches of three consecutive numbers (e.g. 1-3, 4-6, 7-9) and assigned each of the three numbers, in random order, to one of three game variants (baseline, random monitor and elected monitor). This procedure ensured that each variant had an equal number of players in terms of arrival time. This was done in order to prevent systematic biases due to the time of arrival, in case arriving early/late to the activity reflects unobserved heterogeneity that is correlated with the outcomes of the behavioral games. In addition, the number of players in each variant was set to be no fewer than 8, and no larger than 12. In three locations where fewer than 24 sampled members arrived on time, we only played two variants.

4 Interviewer Manual

[These are instructions given to the interviewers during training in addition to the script of the behavioral game.]

Framework:
This game will be played with about 36 Producer Organization (PO) members per Farmer Association. The 36 PO members will be split up into three even groups, with 8-12 PO members in each (Note that the instructions are for situations with 12 players). We will mix PO members to maximize the heterogeneity of the group. Each group will go to a different location, and two interviewers will be responsible for each group. Each group will play one of three variants: The first group will play a baseline public goods game (with no monitor), the second will play a public goods game with a randomly selected monitor, and the third group will play a public goods game with an elected monitor. The baseline public goods game will have six rounds. The variants will each have two "traditional" preliminary rounds with no monitors, followed by four rounds with monitors. Given our focus on the effects of monitor selection, players are assigned to the same group for all six rounds.

Games Summary:
In Public Good Games, players must decide how much of their endowment to keep in their private pocket and how much to put in the group pot. Each player, for each round of the game, receives an endowment of ten 100USH coins totaling to 1000USH. The total amount donated to the group pot is then doubled and redistributed evenly among the players. Payoffs will be rounded up to the nearest 50USH. The contributions will be made public to all the players, but the identity of the contributor will not. The total donation, as well as the average donation, and some exemplary individual payoffs will also be made known to the players.

In the monitor variant there will be a monitor who will be able to spend 100USH to reduce another players' payout by 300USH. The monitor will choose which contributions he wants to reduce. The players will remain anonymous to everyone but the interviewer. In the randomly-selected monitor variant, the monitor will be chosen at random. In the elected variant, the monitor will be elected by the group.

Player i's payoff in round t is calculated as $\pi_{it} = (10 - x_{it}) + \frac{2\sum_{i=1}^{N} x_{it}}{N} - 3P_{it}$, where x_{it} is i's contribution to the public account in round t, and $P_i \in \{0, 1\}$ indicates whether i was punished in round t, and N is the number of players.

Logic:
On an individual level, it is more profitable for a person to "free-ride", and have everyone else contribute to the collective fund. In contrast, the most profitable outcome for the group is if all players contribute all their tokens to the public fund. This game examines how players will balance their own self-interest and the well-being of the collective.

We will be looking at two factors: the effect of monitoring, and the effect of election. We expect that the potential reduction of payoff by the monitor will increase the total contributions to the collective fund by the group. We expect that the election of the monitor will have an

even larger effect on players, since players may be more responsive to monitoring by someone they have elected. We are also interesting in observing monitors' willingness or sense of duty to sacrifice money to sanction someone who donated little or nothing to the collective.

Materials (Per Group):
– 2 large sheets of paper and tape
– 12 3-sided cardboard screens
– 12 sets of ten 100USH
– 12 "Personal Pocket/Collective Pot" boards
– 2 markers
– cards (1-12) for Group 2 only
– calculators
– voting materials for Group 3 only

Setup:
For each Group, there will be two interviewers. The two interviewers will be assigned to be "interviewer 1" and "interviewer 2", as the script will have the interviewers doing different things. Interviewer 1 will talk more, while interviewer 2 will be more focused on data collection and presentation during the game. Interviewer 2 should be prepared to do more math.

Each room/area should have 12 cardboard privacy screens set up with a "Personal Pocket/ Collective Pot" board and ten 100USH coins on the "Personal Pocket" side. There also should be a demonstration set at the front of the room and two large sheets of paper (to serve as a chalkboard). The ideal set up is a circle, but if this is not possible, have the participants sit in rows.

Basic Explanation of the Procedures of the Game:
There are some examples to run through before the start of the first round. These are thorough and each demonstrate one full round. The format of data presentation on the large sheet of paper at the front of the room will be the same for each example and each round.

Here is Example 1: The first series of digits represent how much each individual in the group has donated to the group pot. This has to be reported in increasing order. This data will be recorded by interviewer 2 walking around the room to each player's screen. This list of numbers is very important because it allows players to all see how much everyone in the room donated to the collective, but also allows everyone to remain anonymous.

	Average	Total	Double	Payback
0, 0, 0, 0, 0, 0, 2, 2, 2, 2, 2, 2	100	1200	2400	**200 USH**

It will be important to go through the math of all of them. The "Average" is the sum of the list of numbers and divided by the number of players (12 in rounds without a monitor, 11 in rounds with a monitor). This allows players to see how much each player in the room contributed on average.

The "Total" is the sum of the list of numbers and the "Double" is that sum doubled. "Payback" is the "Double" divided by the number of players in the room (12 in rounds without a monitor, 11 in rounds with a monitor). This ensures that the collective good is now redistributed evenly to each player in the room.

The examples will make clear that: 1) people that donate more to the group pot will make less than people who do not donate or donate fewer coins; 2) the group as a whole gets the most money if everyone puts in all ten coins; 3) if a player puts a lot and everyone else puts in little or nothing, the player can end up with less than the 1000 shillings s/he started out with.

Rounds:
Group 1, playing the Baseline (no monitor) version, will simply repeat the rounds six times. Groups 2 and 3 playing with monitors, will complete two rounds as Baseline, and then 4 rounds according to the monitor variant. The players should not be told how many rounds they are going to play, because there is a tendency to contribute very little when people know they are playing the "final" round.

For Groups 2 and 3, after having played two traditional rounds, the interviewer will inform the players that they will play the game again, but a little differently than in the previous rounds. A monitor will be selected after the first round for the rest of the game. For Group 2, this will be random. For Group 3, the monitor will be elected by the room through written ballots. In both cases, one of the players will become a "monitor". The monitor will not be participating as a regular player. He would therefore not benefit from the decisions of the players and his decision will be simply affected by whether he approves or not the behavior of the players.

The monitor will receive 1000 USH and given the option to give-up 100USH to take away 300USH from contributors whose donation s/he is dissatisfied with. The monitor may do this to as many contributors as he would like, but every single contributor can be penalized only once. The interviewer must also emphasize that the choice is up to the monitor; s/he does not have to reduce anyone's payoff. The interviewer will demonstrate two examples and will ask the monitor a question to ensure comprehension.

After the round is played and all the contributions have been reported on the board, the monitor will have to give the interviewer 100USH from his own endowment for each contributor he wants to penalize and then walk up to the board and mark the numbers corresponding to the contributions s/he wants to "reduce" by drawing an X below the numbers. The identity of the contributors will remain anonymous, but all the players will know the amount of donations that the monitor considers not satisfactory.

Payoff:
At the end of the last round, the interviewer will ask the player sitting at the front right corner to pick a card from a deck numbered 1-6. This determines which round the group will be paid for at the end of the day.

5 Script of the Public Goods Game: Elected Monitor Condition

[Each condition had its own script. The baseline and random monitor conditions are available upon request.]

Game introduction and instructions:
(Interviewer 1 begins instructions, interviewer 2 records everyone's ID numbers on the record.) Hello my name is [interviewer 1] and my colleague's name is [interviewer 2]. Please take a seat at one of the stations. We would like to thank you all for being cooperative and for participating in the various activities.

For this activity, there must be absolute silence. You are not allowed to talk to each other. While in the group, you cannot ask questions or talk. This is very important. Please be sure that you obey this rule, because it is possible for one person to spoil the activity for everyone. If one person talks about the activity while sitting in the group or with other people later, we will not be able to continue the activity. Do not worry if you do not completely understand everything as we go through the examples here in the group. We will take questions when we are finished explaining.

This activity will have a few rounds of decisions. However, only one round will count for payment, which will be chosen randomly once the activity is completed. For the round that is chosen, the money will be yours to take home and use as you please. Since we do not know which round will count, you should decide in each round as if you were deciding on real money.

In front of you there is a board. One side represents your Personal Pocket, the other side represents a Group Pot. There are now 1,000USH in your personal pocket. In this activity, you will have to decide how many Shillings you would like to keep for yourself in your personal pocket, and how many you would like to contribute to the group pot.

You must understand something very important about the group pot. The group pot will include only the contributions from all the people participating in this activity. Once everyone has decided how much to give to the group pot, I will add up the total amount and the research team will double the amount. The group pot will then have twice the amount of money people contributed to it. I will then redistribute the total earnings equally among all twelve of the people participating in the activity. Each person's payoff will be rounded up to the nearest 50USH.

Everyone's donations and decisions will be anonymous, and the screens are here to ensure that. No one will know another's contributions. Do not look at other people's boards. Only the interviewer will know how much each person has donated and the interviewer will never tell anyone else. The number of coins contributed by each person will be reported once all the contributions have been made, but I will not say who donated each number of coins. Names or ID numbers of the people here will not be used throughout the activity.

Let's go over a few examples.
Interviewer 1 explains and demonstrates the coins. Interviewer 2 writes results on the large sheet of paper. Participants should be able to see all three examples at one time by the end of the explanations.
Interviewer 1: One way the activity might turn out is that six participants decide to give nothing and the other six decide to give 200 USH to the group pot. Demonstrate by moving two coins

from Personal Pocket to Group Pot.
Interviewer 2 writes on large sheet of paper and explain each number: "0,0,0,0,0,0,200,200,200,200,200,200", total (1200), average (100), doubling (2400), and Payback (200).

	Average	Total	Double	Payback
0, 0, 0, 0, 0, 0, 200, 200, 200, 200, 200, 200	100	1200	2400	**200 USH**

Interviewer 1: In this case, the results will look like this. We will have a total 1,200 USH in the group pot. The average donation is 100 USH. The research team double the total contribution, so the group pot now has 2,400 USH which will be divided equally amongst all the participants. In this case, everyone will get 200 USH.

The participants that donated 0 to the group pot will get 1000 shillings from their personal pocket and 200 from the group pot, making 1,200 USH. The participants that donated 200 USH will get 800 USH from their personal pocket and 200 USH from the group pot making 1,000 USH.
Interviewer 2: As interviewer 1 explains the individual payoffs, demonstrate on board like the following:

	Average	Total	Double	Payback
0, 0, 0, 0, 0, 0, 200, 200, 200, 200, 200, 200 1000 + 800 + 200 = 200 = 1200 total 1000 total	100	1200	2400	**200 USH**

Let's try another example.

In this case, everyone gives all 1000 shillings to the group pot.
Interviewer 2 write and explain each number: "1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000," total (12000), average (1000), double (24000), payoff (2000)
Interviewer 1 explain: The total contribution was 12,000 USH. In this case, the average contribution was 1,000 USH. Now the research team will double this amount. The group pot now has 24,000 USH. Everyone will get 2,000 USH. Since everyone donated everything, everyone has 0 in their private pocket, and everyone gets 2,000 USH from the group pot. Everyone has made 2,000 USH.

Let's try a final example, in this case 7 people don't give anything to the group pot, 4 people give 100 USH, and 1 person gives 800 USH. The results will look like this:
Interviewer 2 write: "0, 0, 0, 0, 0, 0, 0, 100, 100, 100, 100, 800," total (1200), average (100), double (2400), payoff (200). Also write out the three different contributions and their total pay off during explanation.
Interviewer 1: The total donation is 1200 USH. The average donation is 100 USH, The research team will then double this number so that the group pot has twice the money, or 2400 USH. The payoff is 200 shillings for each person.

The participants that donated nothing will get 1000 USH from their personal pocket, as well as 200 USH from the group pot. This makes 1,200 USH. While the participants that donated

100 will get 900 from their personal pocket, as well as 200 from the group pot. This makes 1100 USH. The participants that donated 800 USH will have 200 USH from their personal pocket, as well as 200 USH from the group. This makes 400 USH.

Here are some things to remember:
People that donate coins will make less than people who do not donate or donate fewer coins. In the first case, 6 people donated 200 USH and 6 people donated 0. The people that donated 200 made 1000 USH, but the people that donated 0 made 1,200 USH.

The group as a whole gets the most money if everyone puts in all 1000 USH. Remember the second example. Everyone donated everything, and everyone made 2,000 USH.

Also keep in mind that if you put a lot and everyone else puts in little or nothing, you can end up with less than the 1000 USH you started out with. Remember the third example. Some made 1,200 USH, while the person who donated the most made only 400 USH.

Before we begin the activity, does anyone have any questions?
[Interviewers should answer questions, but they should stick to the script as much as possible. Also, they do not have to invite too many questions. They have to keep it short!]

Rounds 1 and 2
Interviewer 1: Now please use the screen we have provided and decide how much you would like to contribute to the group pot. We will go around and record your decisions. Place the amount of coins you would like to contribute to the group pot on the right side of the board, with the picture of the group of people. Remember you do not have to contribute if you do not want to. The coins in the group pot will benefit everyone in the room.
Interviewer 2 goes around and records decision on data sheet. Individual contributions are recorded next to the ID of the contributor. Calculate average, total, double, payoff.
Interviewer 2 writes the results of the donations on the large sheet of paper at the front of the

room by writing all contribution amounts in increasing order.
Interviewer 1: Thank you. Please move all your coins back to your personal pocket. Here are the results of all the donations in the room.
Explain each number. Just to clarify, let's go over how much two participants are going to get.
Pick the third and eight donated amount and go over their payments:
The person that donated X will have Y in his personal pocket, and will get Z from the group pot, totaling Y+Z. The person that donated A to the pot will have B in his personal pocket, and get C from the group pot, making B+C.

[Repeat for round 2]

Interviewer 1: Now we will repeat the activity again.
Interviewer 2 walks around to make sure the setup has been followed.
(. . .)

[After round 2]

Monitor Election:
Interviewer 1: We will now repeat the activity again, but a little differently than the previous round. In this round one of the participants will become a "monitor". The monitor will be elected by all of the participants in this activity.

The monitor will not be participating in this activity like everyone else. He will not donate anything to the group pot, and he will not receive any payment from the group pot. He will be given 1000 USH and he can spend 100USH to take away 300USH from the private pocket of contributors he is dissatisfied with. He may do this to as many participants as he would like but he may only use 100USH for each participant. The monitor does not have to reduce from anyone and may keep all 1000 USH.

The monitor will not know the names of the contributors, only how much they donated. He will see the same results as everyone else. The monitor will only choose who to reduce money from by writing an "X" under the number indicating the size of the contribution. Now, we will elect the monitor.

Each of you has a piece of paper and a pen at your stations. Please write the ID number of the person that you would like to elect to be your monitor. An interviewer will walk around to collect your ballots and will tally the results. Please take a moment to look around at the ID tags on the other participants, so you know who you can vote for.

[Interviewer 2 collects ballots, and tallies results, and tells interviewer 1]

Interviewer 1 to elected monitor: You have been elected by the group as the monitor. Please step to the front of the room, and stand beside the large paper. Starting from the next round of activity, you will be able to reduce 300 USH from any contributor, by giving up 100 USH. If you would like to reduce, you will give 100 USH to [interviewer 2], walk up to the board, and mark an X beneath the contributions you would like to reduce shillings from. So if you would like to reduce 300 USH from a person who contributed (point to the second from the left), place an X below it. If you would like to reduce 300 USH from a person who contributed (point to

second from the right), place an X beneath it. Give one coin for each X to [interviewer 2].

Rounds 3-6
Interviewer 1: Now please use the screen we have provided and decide how much you would like to contribute to the group pot. We will go around and record your decisions. Place the amount of coins you would like to contribute to the group pot on the symbol. Remember you do not have to contribute if you do not want to, and the coins in the group pot will benefit everyone in the room.
Interviewer 2 goes around and record decision on data sheet. Individual contributions are recorded next to the ID of the contributor. Calculate average, total, double, payoff.
Interviewer 1 writes the results of the donations on the large sheet of paper at the front of the room by writing all contribution amounts in increasing order.
Interviewer 1: Please move all your coins back to your personal pocket. This is the result of all the donations in the room.
Let's go over how much two participants are going to get.

Pick the third and eight donated amount and go over their payments:
The person that donated X will have Y in his personal pocket, and will get Z from the group pot, totaling Y+Z. The person that donated A to the pot will have B in his personal pocket, and get C from the group pot, making B+C. Now, the monitor can decide if he would like to reduce 300 USH from a contributor.
Interviewer 1 to monitor: Would you like to reduce the 300 USH from any contributor? If yes, please mark an X by the ones you want to reduce from and give 100 USH to [interviewer 2] for each X you mark. If not, let us know now.
Explain results to the room The Xs mean that 300 USH have been reduced from the person that donated the amount indicated to the group pot. So for example, one person that donated X has just had 300 USH reduced from their payoff. Instead of earning X, they will now earn Y.
Announce repetition Now, we will repeat the activity again. After you've made a decision the elected monitor will again be given the possibility to take some money away from you. Please move your coins back to their original position so that there are ten coins in your private pot.
(. . .)
[This process should be repeated for a total of 4 rounds.]

Payoff (end of 6 rounds):
The cards are numbered from one to six. We will ask one of you to pick a card. The number that will be picked will determine which round you will be paid for at the end of the day. For example, if you pick card number one you will be paid for round 1 (point to round 1 results). If you pick card number three you will be paid for round 3 (point to round 3 results). Remember, you will get both what you kept in your private pocket, and the payoff from the group pot.

Interviewer 1 asks the front right participant: Please pick a card from this shuffled deck.

Interviewer 2 records accordingly.
Thank you for your time. You may now leave.