Online Appendix for: Identifying communication spillovers in lab-in-the-field experiments

Alexander Coutts*

November 2020

A Results for Distance of 2 km

This section replicates the central results of the analysis in the main paper for the distance of 2 km rather than 1.75 km. Results are broadly consistent, across both distances. There are two differences worth noting for the 2 km distance cutoff. First, the results on conditional cooperators are similar, although no longer significant at conventional values (p-value on interaction term in column 3 of Table A5 is 0.110). Second, regarding heterogeneous effects by distance to nearest neighbors in Table A4, the effect of short distances is now statistically significant in column 3, at the 5% level.

Table A1: Balance of treatment: Past participating neighbors within 2 km

	Treatment	Control	Difference
Available to planner			
Distance to base (km)	11.92	15.02	-3.10**
Village Size (# HHs)	131.24	132.24	-1.00
# Villages $\leq$ 2 km	3.38	1.84	1.54***
Distance to paved road (km)	4.60	4.93	-0.33
Unavailable to planner			
Contribution	267.00	237.56	29.44***
Female	0.74	0.73	0.01
Age	35.25	35.59	-0.35
Years of Education	4.50	4.42	0.08
Community Cooperation Index	0.82	0.82	-0.00
Community Effort Index	0.79	0.79	0.00
Number of Strong Ties	2.52	2.63	-0.11
General Trust Index	0.73	0.73	-0.00
Observations	89	58	147

Significant differences indicated by * 0.1; ** 0.05; *** 0.01.

*Nova School of Business and Economics, Universidade Nova de Lisboa, Campus de Carcavelos, Rua da Holanda 1, 2775-405 Carcavelos, Portugal, NOVAFRICA, and Schulich School of Business, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3; alexander.coutts@novasbe.pt

Table A2: ATE of Presence of Past Participating Neighbors Within 2 km

	(1) Standard Matching	(2) Exact Matching
Contribution	36.143***	36.177**
	(11.248)	(14.823)
Observations	113	89

Analysis uses nearest neighbor propensity score matching, with 2 neighbors, with replacement. Significantly different from zero at * 0.1; ** 0.05; *** 0.01. Abadie-Imbens Robust Standard Errors in parentheses. Values of propensity score outside common support range are dropped. Exact matching excludes sectors with only 0 or 1 village in either treatment or control groups.

Table A3: Effect of Presence of Past Participating Neighbors Within 2 km

	(1)	(2)	(3)
Treatment Status	29.440***	32.084***	30.868**
	(10.410)	(10.740)	(12.883)
Distance to base (km)		0.000	−0.092
		(0.844)	(3.223)
Distance to paved road (km)		4.708**	1.903
		(1.738)	(4.524)
Village Size (# HHs)		−0.020***	0.062***
		(0.110)	(0.136)
# Villages $\leq$ 2 km		−1.890	0.246
		(4.405)	(5.197)
Years of Education		3.999	2.282
		(4.747)	(6.219)
Female		99.687***	89.418**
		(36.015)	(41.798)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.05	0.20	0.29
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in main paper.

Table A4: Effect of Presence of Past Participating Neighbors Within 2 km (By time elapsed)

	(1)	(2)	(3)
Treatment Status (short)	34.554**	36.976**	40.089**
	(13.752)	(15.060)	(17.216)
Treatment Status (medium)	31.005**	33.221**	32.229*
	(14.298)	(14.049)	(16.461)
Treatment Status (long)	21.207	23.702	17.491
	(15.377)	(15.453)	(19.379)
Distance to base (km)		−0.057	−0.028
		(0.885)	(3.171)
Distance to paved road (km)		4.743***	2.196
		(1.727)	(4.385)
Village Size (# HHs)		−0.031	0.050
		(0.113)	(0.139)
# Villages $\leq$ 2 km		−1.901	0.324
		(4.470)	(5.312)
Years of Education		4.152	2.064
		(4.686)	(6.281)
Female		103.481***	90.952**
		(36.735)	(42.080)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.05	0.21	0.30
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in main paper. Short refers to the average date of neighbors' participation was < 3 days, medium refers to $3 - 7$ days, and long refers to > 7 days.

Table A5: The Role of Conditional Cooperators (CCs)

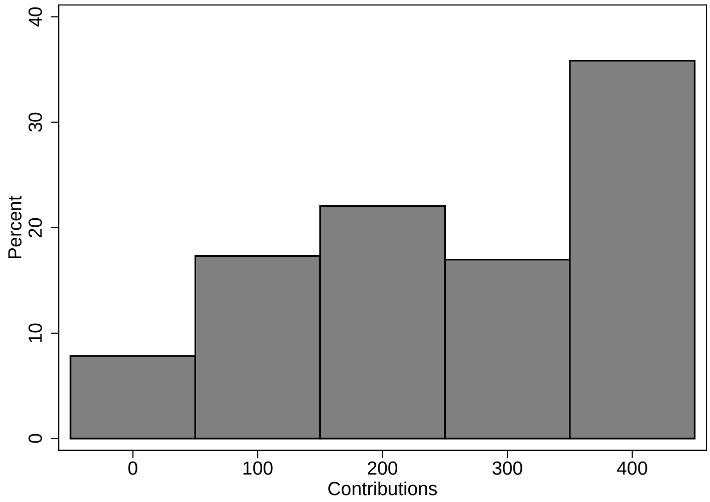
	(1)	(2)	(3)
Treatment Status	2.286	2.523	−4.537
	(12.244)	(13.050)	(15.514)
Conditional Cooperator (CC)	119.870***	103.054***	100.735***
	(31.993)	(32.611)	(31.623)
CC × Treatment	45.815	50.360	58.707
	(35.496)	(35.880)	(36.419)
Distance to base (km)		−0.130*	−2.681
		(0.649)	(2.572)
Distance to paved road (km)		2.243	1.580
		(1.322)	(3.787)
Village Size (# HHs)		−0.044***	0.046***
		(0.098)	(0.111)
# Villages ≤ 2 km		−1.131	1.326
		(3.096)	(3.423)
Years of Education		1.386	0.645
		(4.066)	(4.727)
Female		55.768*	33.425
		(28.377)	(32.518)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.44	0.50	0.58
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in the main paper. CC refers to the proportion of individuals classified as conditional cooperators within each village.

B Individual Contributions

Figure B1 presents the distribution of individual contributions in the public goods games at the individual level. The possible levels ranged from 0 to 400 RWF, in 100 RWF increments. Only 8% of individuals contribute the self-interested profit maximizing choice of 0, while the most common contribution is the maximum of 400, sent by 36% of individuals. Note that due to 14 missing observations, $N = 1750$, across 147 villages.

Figure B1: Distribution of Individual Contributions (RWF)



Individual contributions in the public goods experiments. Possible values ranged from 0 to 400 RWF, in 100 RWF increments. N = 1750.

C Supplementary Variables for Additional Balance Tests

This subsection presents additional balance on household characteristics, aggregated to the village level, which would be unobserved by the planner. These variables come from a different baseline survey of the evaluation of community health clubs, which was conducted before the public goods games. One village was not able to be matched. Due to the nature of that evaluation, many of the variables are focused on health, while there are a few which relate to household assets and construction quality. There were no questions on religion nor on ethnicity (not permitted by Rwandan law). From Table C1 one can see that from this small set of variables, there are no significant differences across the villages assigned to treatment versus control.

Table C1: Balance of treatment: Past participating neighbors within 1.75 km (Additional Variables)

	Treatment	Control	Difference
Unavailable to planner			
Anyone with diarrhea < 7 days	0.17	0.18	-0.01
Anyone ill or injured < 4 weeks	0.60	0.60	-0.00
Number of livestock	2.34	2.46	-0.11
Durable asset index	0.77	0.73	0.04
Anyone didn't sleep under bed net	0.12	0.14	-0.02
Log of value of house	12.97	12.84	0.13
Housing materials index	0.67	0.65	0.02
Anyone take medication for fever < 4 weeks	0.26	0.27	-0.01
Observations	73	73	146

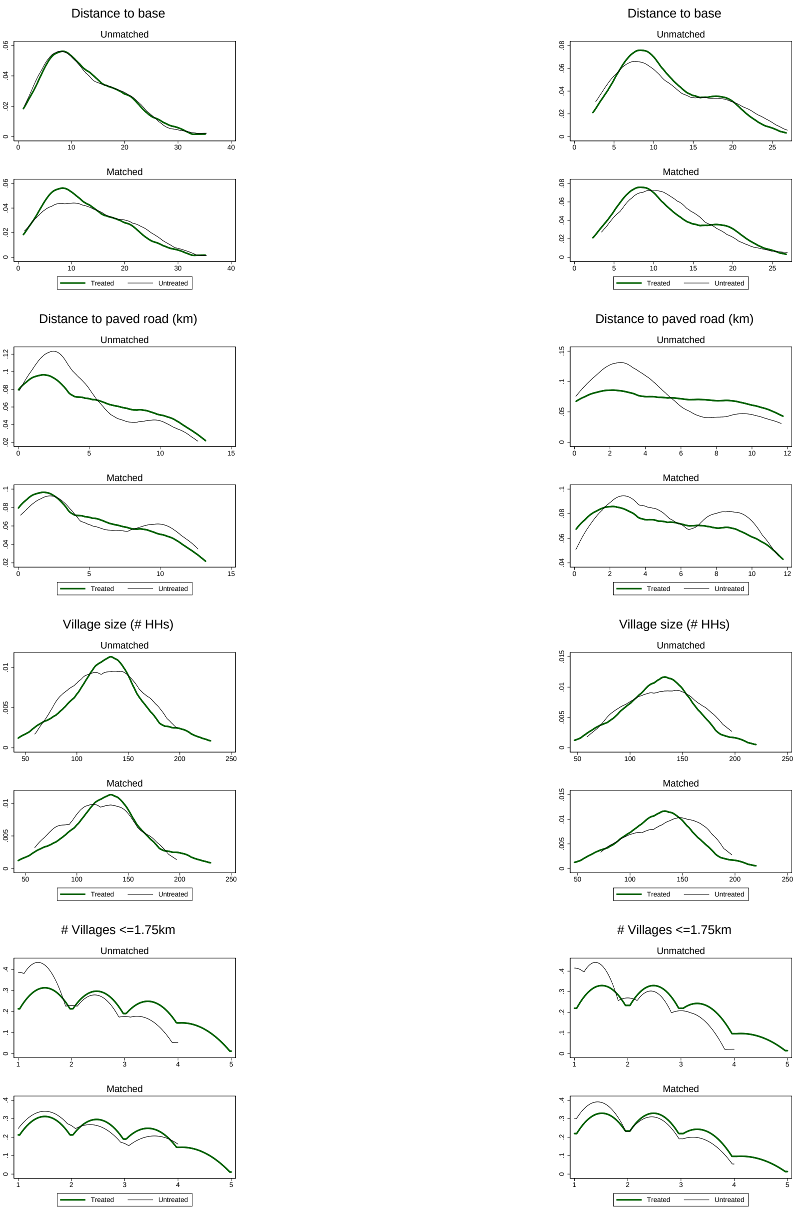
Significant differences indicated by * 0.1; ** 0.05; *** 0.01.

D Balance of Matching Variables

Figure D1 shows the distributions of matching variables by treatment, in order to evaluate balance of the matching strategy. There are not large imbalances across these variables, even in the unmatched stage. Matching improves balance, which is slightly better in standard matching (column (a)) rather than exact matching (column (b)), though the differences are not substantial.

Figure D1: Balance of Matching Variables

(a) Standard Matching (b) Exact Matching



Density of covariates in matching estimations by treatment, before and after matching. **(a)** $N = 118$ **(b)** $N = 102$.

E Full OLS Specification

Table E1 presents Table 3 in the main paper with the full set of controls shown. Of note is that the three survey variables on levels of trust, cooperation, and effort exertion are all significantly correlated with one another, with ρ varying from 0.48 to 0.58. If these variables were to enter individually, the trust and cooperation measures are positive and significant at the 5% and 10% level respectively, in column 3. Note also that the number of strong ties is significantly associated with fewer contributions. I do not have an explanation for this result, although it should be noted that *strong ties* indicates the number of others one knows by both first and last name. Thus these links can be unilateral, and may not be social in nature.

Table E1: Effect of Presence of Past Participating Neighbors Within 1.75 km

	(1)	(2)	(3)
Treatment Status	30.247***	28.375***	27.736**
	(10.517)	(10.214)	(11.623)
Distance to base (km)		−0.072	−1.001
		(0.830)	(3.182)
Distance to paved road (km)		4.354**	0.973
		(1.705)	(4.435)
Village Size (# HHs)		0.010	0.090
		(0.112)	(0.138)
# Villages $\leq$ 1.75 km		0.243	2.665
		(4.711)	(5.612)
Years of Education		2.312	−0.027
		(4.905)	(6.347)
Female		97.033***	86.295**
		(37.274)	(43.021)
Age		12.839	16.726
		(26.332)	(25.826)
Age^2		−0.171	−0.234
		(0.364)	(0.359)
General Trust Index		63.993	75.653
		(48.118)	(48.621)
Community Cooperation Index		53.488	55.765
		(51.655)	(65.730)
Community Effort Index		−20.366	−26.792
		(40.272)	(41.970)
Number of Strong Ties		−11.913**	−9.792*
		(5.420)	(5.859)
Sector Fixed Effects			✓
R^2	0.05	0.20	0.30
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in main paper.

F Robustness to Alternative Variables

F.1 Different Approaches to Measuring Opportunities for Communication

In this section I examine three alternative variables which are likely to be correlated with opportunities for communication with past participants. In the section directly following, I examine two placebo tests with variables which relate to the overall number of villages (either in the sector or in the 1.75 km radius), and show that these are not significantly related to contributions.

Table F1 presents the effect of the order that a village was visited within a sector. That is, a village that is the very first to participate in the public goods games in its sector would be coded as 1, the village that is the second to be visited within its sector would be coded as 2, and so on. Then the variable is standardized to be mean zero, standard deviation one, in order to facilitate comparison across tables in this section.

Villages that participate first in their sector will be very unlikely to have had contact with previous participants. While villages who participate after 5 villages in their sector had participated, will be much more likely to have had such contact. Consistent with this and the results of the main paper, Table F1 shows that the order of visit within sector is associated with significantly higher contributions.

The next Table F2 simply counts the number of villages within the sector which previously participated in the games. Again this variable is positive and significant in determining contributions. Again it has been standardized. As a placebo test, in the next section, Table G1 shows that the total number of villages in the sector is not driving this result.

Finally Table F3 examines the effect of a variable defined as the distance to the nearest past participating village. This variable ranges from 0.2km to 27km, with a mean value of 2.5km. As very few villages have distant neighbors, this variable is top-coded at the 95^{th} percentile, 5.17km.[1] From column 3, one can see that for every additional kilometer to the nearest participating village, contributions are reduced by 9.2, or about 4% of the average.

[1] The reason for top-coding is that one would not expect any relationship between distance and cooperation among the further distances, since these villages are most likely to have had no communication independent of distance. Hence these distances simply introduce noise. Without top-coding, the estimates continue to be significant in column 3, but not in columns 1 and 2.

Table F1: Effect of Ranking of Order Visited Within Sector

	(1)	(2)	(3)
Rank of Visit (within Sector)	16.423***	15.339***	17.291***
	(5.145)	(4.688)	(6.043)
Distance to base (km)		0.214	−1.670
		(0.863)	(2.982)
Distance to paved road (km)		3.810**	−0.303
		(1.750)	(4.382)
Village Size (# HHs)		0.005	0.062
		(0.112)	(0.132)
# Villages $\leq$ 1.75 km		5.345	8.828*
		(4.481)	(4.902)
Years of Education		3.481	3.182
		(4.738)	(6.607)
Female		104.160***	90.462**
		(35.737)	(40.331)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.06	0.21	0.31
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Variable "Rank of Visit (within Sector)" has been standardized. Controls includes all remaining variables found in Table 1 in main paper.

Table F2: Effect of Number of Previous Participating Villages in Sector

	(1)	(2)	(3)
# Villages before in Sector	12.987***	13.938***	15.185**
	(4.953)	(4.886)	(5.911)
Distance to base (km)		0.437	−1.762
		(0.858)	(2.997)
Distance to paved road (km)		4.600***	−0.052
		(1.772)	(4.390)
Village Size (# HHs)		−0.007	0.055
		(0.111)	(0.133)
# Villages $\leq$ 1.75 km		5.991	9.413*
		(4.478)	(4.851)
Years of Education		4.717	2.843
		(4.854)	(6.471)
Female		101.796***	88.682**
		(36.401)	(40.652)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.04	0.21	0.30
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Variable "# Villages before in Sector" has been standardized. Controls includes all remaining variables found in Table 1 in main paper.

Table F3: Effect of Distance to Nearest Past Participating Village

	(1)	(2)	(3)
Distance to Nearest Past Participant	-8.939^{**}	-7.477^{**}	-9.166^{**}
	(3.724)	(3.743)	(4.252)
Distance to base (km)		0.049	-0.659
		(0.837)	(3.208)
Distance to paved road (km)		4.573^{**}	1.787
		(1.789)	(4.872)
Village Size (# HHs)		-0.034	0.039
		(0.116)	(0.143)
# Villages $\leq$ 1.75 km		3.147	6.028
		(4.437)	(5.031)
Years of Education		4.237	1.201
		(5.129)	(6.681)
Female		89.284^{**}	80.349^{*}
		(40.385)	(43.581)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.03	0.17	0.27
Observations	144	144	144

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in main paper. Distance to nearest past participant is in kilometers, and is top-coded at the 95^{th} percentile (5.17km). The 3 missing observations are because the first 3 villages (on day 1) have no past participating neighbors.

F.2 Alternative Definition of Village Density

A necessary condition for unbiased OLS estimates from Table 3 in the main paper is that the treatment variable is plausibly random, conditional on village density. However in this case, the probability of treatment is not a linear function of village dens ity: the change in the probability of being treated increases more as one moves from 0 to 1 neighbors, than if one moves from 4 to 5 neighbors. Hence this mispecification may result in biased OLS estimates. To account for this, Table F4 replaces the village density variable with the conditional probability of treatment, for each village density value from 0 to 5. These probabilities are respectively: 0, 0.44, 0.61, 0.63, 0.82, and 1. From Table F4 one can see that the estimates of treatment status are not substantively affected by this modification.

Table F4: Effect of Presence of Past Participating Neighbors Within 1.75 km

	(1)	(2)	(3)
Treatment Status	30.247***	32.613***	30.464**
	(10.517)	(10.620)	(11.874)
Distance to base (km)		−0.183	−1.059
		(0.826)	(3.209)
Distance to paved road (km)		4.377***	1.260
		(1.700)	(4.463)
Village Size (# HHs)		0.013	0.090
		(0.111)	(0.137)
Prob. Treatment		−16.600	−0.698
		(23.750)	(25.654)
Years of Education		2.705	0.307
		(4.902)	(6.309)
Female		101.501***	89.976**
		(36.870)	(43.385)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.05	0.20	0.29
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in main paper. Prob. Treatment is equal to the empirical probability of treatment, conditional on the value of village density (# villages within 1.75 km).

G Placebo Tests for Measuring Opportunities for Communication

This section presents three placebo tests for measuring opportunities for communication. The previous Table F2 examined the number of villages that were previously visited within the same sector, and found a strong significant relation with contributions. Table G1 examines the total number of villages in the sector, irrespective of the date of visit. This is correlated, but not significantly associated with contributions. Note that this variable is standardized to have mean zero, standard deviation one, to be comparable with the other tables in this section.

The second and third placebo tests are in fact the variable reflecting village density, the number of villages within 1.75 km, irrespective of the sector or date of visit. This variable is a key matching variable, but it is of note that even though it is correlated with treatment, there are not statistically significant effects of village density on contributions, shown in Table G2. A positive coefficient is to be expected, given that the variable is correlated with treatment. It is significantly diminished when the treatment variable is added to the regression (column 4). Again note that the variable has been standardized.

Table G3 presents this same analysis, but restricted only to the control sample. That is, these regressions examine the effect of the total number of neighboring villages within 1.75km, for the case when none of these neighboring villages participated before. Reassuringly, there is no relationship between village density and contributions. Additionally the coefficient is much smaller in magnitude than Table G2, and sometimes negative.

Table G1: Effect of Total Number of Villages Within Sector

	(1)	(2)
# Villages in Sector	5.869	4.973
	(4.620)	(5.126)
Distance to base (km)		0.216
		(0.913)
Distance to paved road (km)		4.720***
		(1.806)
Village Size (# HHs)		−0.002
		(0.117)
# Villages ≤ 1.75 km		5.284
		(4.538)
Years of Education		3.257
		(4.973)
Female		92.697**
		(38.471)
Controls		✓
Sector Fixed Effects		
R^2	0.01	0.17
Observations	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Sector fixed effects omitted as there is no variation across sectors in number of villages within the sector. Variable has been standardized. Controls includes all remaining variables found in Table 1 in main paper.

Table G2: Effect of Number of Neighbors Within 1.75 km

	(1)	(2)	(3)	(4)
# Villages ≤ 1.75 km	8.857	7.567	11.182	3.696
	(5.874)	(6.283)	(6.866)	(7.783)
Treatment				27.736**
				(11.623)
Distance to base (km)		−0.087	−0.835	−1.001
		(0.860)	(3.163)	(3.182)
Distance to paved road (km)		4.782***	0.848	0.973
		(1.811)	(4.424)	(4.435)
Village Size (# HHs)		−0.012	0.061	0.090
		(0.120)	(0.139)	(0.138)
Years of Education		2.653	0.121	−0.027
		(4.919)	(6.299)	(6.347)
Female		93.179**	74.891*	86.295**
		(38.157)	(40.966)	(43.021)
Controls		✓	✓	✓
Sector Fixed Effects			✓	✓
R^2	0.02	0.17	0.27	0.30
Observations	147	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Variable "# Villages ≤ 1.75 km" has been standardized. Controls includes all remaining variables found in Table 1 in main paper.

Table G3: Effect of Number of Neighbors Within 1.75 km (Control Only)

	(1)	(2)	(3)
# Villages $\leq$ 1.75 km	1.585	-2.327	3.318
	(6.174)	(7.172)	(8.529)
Distance to base (km)		-0.943	-6.581^{*}
		(1.007)	(3.882)
Distance to paved road (km)		3.357	0.785
		(2.339)	(5.161)
Village Size (# HHs)		-0.009	0.162
		(0.119)	(0.160)
Years of Education		1.002	-4.189
		(5.767)	(7.399)
Female		96.928^{*}	63.157
		(53.756)	(69.610)
Age		-8.572	-6.640
		(36.322)	(47.821)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.00	0.19	0.47
Observations	74	74	74

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Variable "# Villages $\leq$ 1.75 km" has been standardized. Controls includes all remaining variables found in Table 1 in main paper.

H Communication Spillovers Across Treatment Versions

This section presents more detailed analysis to Section 2.8 in the main paper. The primary evidence shows that communication spillovers can bias the measurement of preferences in a single version of a public goods game which was played by all villages in the first round. However, many lab-in-the-field games examine the impact of different treatments, by playing different versions with different groups of participants. A possibility is that communication spillovers could impact different treatment versions equally, and thus not lead to bias in estimation of the resulting 'treatment effects'. To investigate whether this may be the case, I make use of the second round of public goods games which utilized one of four different versions, which were selected randomly.

The first version, played in 52% of villages, was the standard baseline game repeated. Note that this version was intentionally over-sampled. The second version, played in 17% of villages, involved punishment. Specifically, after making contributions, each player was anonymously matched with another player whose contribution they saw, and could decide whether to give a penalty. The penalty would remove 100 RWF from the earnings of the matched player, and would cost 50 RWF to the punishing player. The third version, played in 15% of villages, was an analogous reward version, where players could add 100 RWF to a matched player, at a cost of 50 RWF. The fourth version, played in 16% of villages, involved aggregate risk. Rather than having the group fund multiplied by a factor of 3, individuals were given two coin purses and told there were two funds: after all contributions had been made in either (or both) purses, a coin would be flipped to determine which fund would be multiplied by a factor of 5, with the other fund remaining constant. This introduced aggregate risk, although the expected value of contributing was identical. More details can be found in participant instructions in Section N.1.

Given that there are few villages playing the three alternate versions, ranging from 22 to 25, I first pool these versions together, defining the dummy variable *Alt. Version*. I then estimate whether the impact of the alternate versions, compared to the baseline version, was different depending on whether a village had past participating neighbors within 1.75 km (i.e. treatment status). To do so I examine the same OLS specification found in Table 3 in the main paper, with the addition of the alternate version variable and its interaction with treatment status.

Table H1 presents this specification. Consistent with Figure 3 in the main paper, the impact of the alternate version is starkly different depending on treatment status. Specifically, while the alternate version has a large and significant positive effect in villages which were unlikely to have had communication spillovers (control), the interaction with treatment status is equally large and significant, but negative. Thus, while control villages significantly increased their contributions in alternate versions of the game relative to the baseline, there was no impact in treatment villages. As treatment status is positive and significant, this shows that treatment villages which had past participating neighbors (likely affected by communication spillovers) had overall elevated levels of contributions, independent of whether the version was baseline or not. By contrast, control villages had lower levels of contributions for the baseline version, which substantially increased in alternate versions.

Overall these results clearly show that communication spillovers can impact estimated treatment effects across different versions. Although not shown here, if one was not aware of spillovers and did not account for them, the estimated overall impact of the alternate versions in the column 3 specification would be 19 RWF, and not significant (p-value 0.215). Thus, although the true hypothetical impact among a set of villages which were not subject to communication spillovers appears to be large and positive, ignorance of these spillovers would lead one to erroneously conclude that there is no effect.

Finally for completeness, Table H2 presents the analogous specification separating alternate versions into the three possible versions: penalty, reward, and aggregate risk. One can see that the patterns are similar, although not always significant.

Table H1: Effect of Presence of Past Participating Neighbors Within 1.75 km (Second Round)

	(1)	(2)	(3)
Treatment Status	54.378***	47.057***	38.907**
	(16.659)	(16.644)	(18.860)
Alt. Version	50.104***	48.951***	44.484**
	(13.956)	(17.380)	(19.072)
Alt. Version × Treatment	−61.274***	−56.867***	−50.565**
	(20.890)	(21.548)	(23.331)
Distance to base (km)		0.419	2.383
		(0.867)	(3.153)
Distance to paved road (km)		3.485*	0.703
		(1.840)	(4.074)
Village Size (# HHs)		−0.054	−0.071
		(0.115)	(0.137)
# Villages ≤ 1.75 km		3.208	4.722
		(4.735)	(5.203)
Years of Education		3.018	3.178
		(5.137)	(6.425)
Female		65.810*	47.863
		(38.412)	(44.301)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.11	0.23	0.30
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is second round contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in main paper. Alt. Version is indicator for whether second round was one of three alternate versions: reward, penalty, or aggregate risk.

Table H2: Effect of Presence of Past Participating Neighbors Within 1.75 km (Second Round)

	(1)	(2)	(3)
Treatment Status	54.378***	47.974***	39.713**
	(16.897)	(16.814)	(19.283)
Penalty Version	59.838***	55.445***	50.182**
	(18.549)	(21.379)	(23.086)
Penalty × Treatment	−68.524***	−65.904**	−58.404*
	(26.324)	(28.326)	(29.903)
Reward Version	65.073***	71.353***	65.160***
	(16.660)	(22.780)	(23.432)
Reward × Treatment	−56.297*	−54.855*	−43.373
	(30.837)	(33.338)	(35.863)
Risk Version	22.270	21.851	18.641
	(19.727)	(20.133)	(25.156)
Risk × Treatment	−51.076**	−43.332*	−43.097
	(25.250)	(24.583)	(28.754)
Distance to base (km)		0.358	2.089
		(0.891)	(3.189)
Distance to paved road (km)		3.895**	1.588
		(1.873)	(4.208)
Village Size (# HHs)		−0.033	−0.042
		(0.113)	(0.137)
# Villages ≤ 1.75 km		2.379	3.506
		(4.892)	(5.427)
Years of Education		3.603	4.112
		(5.280)	(6.530)
Female		55.528	34.610
		(37.427)	(44.503)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.14	0.26	0.33
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is second round contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in main paper. Reward, penalty, and risk refer to the different second round game versions – see text.

I Instrumental Variable Analysis

Here I examine an instrumental variables strategy. The aim of this exercise is to find a set of instruments which are correlated with the treatment variable, but are otherwise exogenous in relation to contributions in the public goods game. To find such instruments, I simulate six hypothetical routes, using the three observables available to the planner, in increasing and decreasing order (as in Table 5 in the main paper). The hypothetical ordering of villages is conducted as follows. I first rank the 18 political sectors along the dimensions of these three variables (low to high, and vice-versa), taking the average value of the variable per sector. I mandate that the order of visit follow this sector ranking, such that a sector is visited in its entirety, before proceeding to the next sector in the ranking. Within each sector, the order of village visits follows the ranking of the variable under consideration.[2] The next step is to define the number of hypothetically treated neighbors (within 1.75 km), which is the instrument. Because the actual route was chosen principally based on logistical convenience (making use of these observables), and proceeded largely sector by sector, it is reasonable to hypothesize that this instrument will be correlated with the treatment derived from the actual route.

This process generates six potential instruments. To select a set of "optimal" instruments, I follow the post-double selection method of Belloni et al. (2012), which makes use of methods for sparsity, Lasso and Post-Lasso, to select the optimal instruments. The results of these methods select three instruments: (1) the number of participating neighbors generated by the hypothetical ordering which orders sectors (and then villages within sectors) according to their distance to a paved road (from low to high); (2) and (3) the analogous instruments which derive from an ordering by distance to the team's base location, from low to high and from high to low.[3]

Table I1 presents the 2SLS regressions using identical covariates as the primary OLS analysis, Table 3, in the main paper. The estimated coefficients are statistically significant at the 5% level, and significantly larger in magnitude than the matching or OLS estimates. They suggest an effect of communication of about 51 RWF, or a 20% increase in contributions. One note of caution is that these estimates are less precisely estimated, as the standard errors are substantially larger. I cannot reject that the 2SLS estimate is equal to the corresponding OLS coefficient in the main paper, for column (3).[4]

[2]Without imposing such sequentiality, the resulting number of hypothetically treated neighbors is uncorrelated with the treatment, invalidating its use as an instrument. To give an example, one variable is distance to the nearest paved road (ranked low to high). The first sector visited would be the sector with the lowest average distance to the nearest paved road. The first village to be visited would be the village within that sector nearest to a paved road. The following villages within that sector would be visited according to how far they are from a paved road, from shortest to longest. After all villages in the first sector are visited, the second sector chosen is the one with the second lowest average distance to the nearest paved road. This process continues until all villages are visited.

[3]In order to not violate the exclusion restriction, these two variables: distance to a paved road and distance to base must be included in the 2SLS regression.

[4]A final important note is that one explanatory variable, the density within 1.75 km is replaced by the density within 3 km, due to collinearity with the derived instrumental variables. Without this modification the results (available on request) are consistently large, with coefficients on treatment status across columns (1) to (3) equaling approximately 49, 65, and 91. However, the instruments are too weak to make any serious inference, and the latter two coefficients are not significant at conventional levels.

Table I1: IV-2SLS - Effect of Presence of Past Participating Neighbors Within 1.75 km

	(1)	(2)	(3)
Treatment Status	49.478*	52.279**	50.611**
	(25.582)	(24.822)	(21.585)
Distance to base (km)	1.951	−1.039	−1.966
	(1.660)	(0.943)	(2.988)
Distance to paved road (km)	0.113	4.538***	1.910
	(0.786)	(1.554)	(4.108)
Village Size (# HHs)		0.041	0.121
		(0.108)	(0.125)
# Villages $\leq$ 3 km		−4.539	−2.668
		(3.004)	(3.195)
Years of Education		1.812	−0.224
		(4.806)	(5.761)
Female		99.634***	96.064**
		(35.443)	(39.378)
Controls		✓	✓
Sector Fixed Effects			✓
First stage F-Statistic	21.12	12.83	11.94
R^2	0.05	0.19	0.28
Observations	147	147	147

Analysis uses instrumental variable two-stage least squares (IV-2SLS) regression. Dependent variable is contributions. See text for instrument selection. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in the main paper.

J Examining Potential Changes in Cooperation Norms

Table J1 examines whether the treatment, having past participating neighbors within 1.75 km, had any impact on the response to the question, "People in this community generally cooperate with one another on issues that affect the community". This question was asked before the public goods game was introduced. From this table it is evident that responses to this question were not affected, providing some evidence that community norms about cooperation were not significantly altered. Given the main results in the paper, it appears more plausible that communication changed beliefs specific to behavior in the game, rather than norms in general.

Table J1: Effect of Presence of Past Participating Neighbors Within 1.75 km on Cooperation Norms

	(1)	(2)	(3)
Treatment Status	−0.013	−0.016	−0.008
	(0.022)	(0.021)	(0.019)
Distance to base (km)		−0.001	0.002
		(0.001)	(0.006)
Distance to paved road (km)		0.000	−0.007
		(0.002)	(0.007)
Village Size (# HHs)		0.000	0.000
		(0.000)	(0.000)
# Villages $\leq$ 1.75 km		0.013	0.014*
		(0.008)	(0.008)
Years of Education		−0.006	0.003
		(0.008)	(0.009)
Female		0.013	0.110*
		(0.062)	(0.066)
Age		0.057*	0.062*
		(0.033)	(0.034)
Age^2		−0.001*	−0.001**
		(0.000)	(0.000)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.00	0.44	0.58
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in main paper.

K Restricting OLS Estimation of Conditional Cooperation

In this section Table K1 presents the analog to Table 6 in the main paper, but restricting the analysis to only villages which repeated the identical public goods game when playing it for a second (final) round. This lowers the number of villages to 76, 52% of the sample. The other villages played one of three different versions involving opportunities to reward or punish, or a public goods game involving risk. These versions were randomly selected. The villages only knew about the existence of their particular version after having played the first round. From Table K1 it is possible to verify that the results are consistent with the analysis which involves the whole sample. This provides some reassurance that the results in Table 8 are not driven by changes in the versions of the game played.

Table K1: The Role of Conditional Cooperators (CCs)

	(1)	(2)	(3)
Treatment Status	−13.680	−19.087	−34.695*
	(18.308)	(17.455)	(20.207)
Conditional Cooperator (CC)	106.261**	86.562*	124.566***
	(46.522)	(52.407)	(45.937)
CC × Treatment	93.630*	101.597*	93.871*
	(49.662)	(56.838)	(53.679)
Distance to base (km)		−0.747	−8.716***
		(0.797)	(3.384)
Distance to paved road (km)		1.772	3.721
		(1.609)	(5.242)
Village Size (# HHs)		−0.097	−0.126
		(0.124)	(0.123)
# Villages ≤ 1.75 km		0.962	2.278
		(6.109)	(7.061)
Years of Education		−5.223	−11.296
		(5.825)	(8.477)
Female		20.576	3.227
		(46.614)	(48.892)
Age		9.893	12.135
		(25.415)	(27.662)
Age^2		−0.183	−0.209
		(0.353)	(0.381)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.52	0.64	0.79
Observations	76	76	76

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in the main paper. CC refers to the proportion of individuals classified as conditional cooperators within each village.

L Heterogeneous Effects with Matching Strategy

Table L1 presents the analogous examination to Appendix Table 14 in the main paper, using the matching strategy. The exercise is fairly demanding, since it requires matching on very small subsamples. One can see that for standard matching there are still a moderate amount of observations (between 57 and 60) allowing for some inference - in particular the patterns in column (1) are similar to those in Table 14. However exact matching by sector is far too demanding, and there are not sufficient observations to be matched, ranging from 20 to 22. Thus the results in Column (2) are presented only for consistency, it is not possible to draw any conclusions from this data.

Table L1: Average Effect: Heterogeneous Effects

	(1) Standard Matching	(2) Exact Matching
By Conditionally Cooperative		
More Conditionally Cooperative	45.918***	5.579
	(15.295)	(15.935)
Less Conditionally Cooperative	12.897	−3.485
	(10.772)	(9.165)
N	60	20
By Age		
Older	41.095***	30.283*
	(14.269)	(15.916)
Younger	8.492	11.913
	(15.939)	(25.699)
N	57	20
By Education		
More Educated	28.446*	79.412***
	(14.601)	(13.639)
Less Educated	36.643**	31.875
	(14.895)	(33.619)
N	57	20
By Strong Ties		
More Strong Ties	25.625*	42.803
	(15.132)	(31.578)
Fewer Strong Ties	27.788*	5.975
	(15.563)	(37.759)
N	58	22

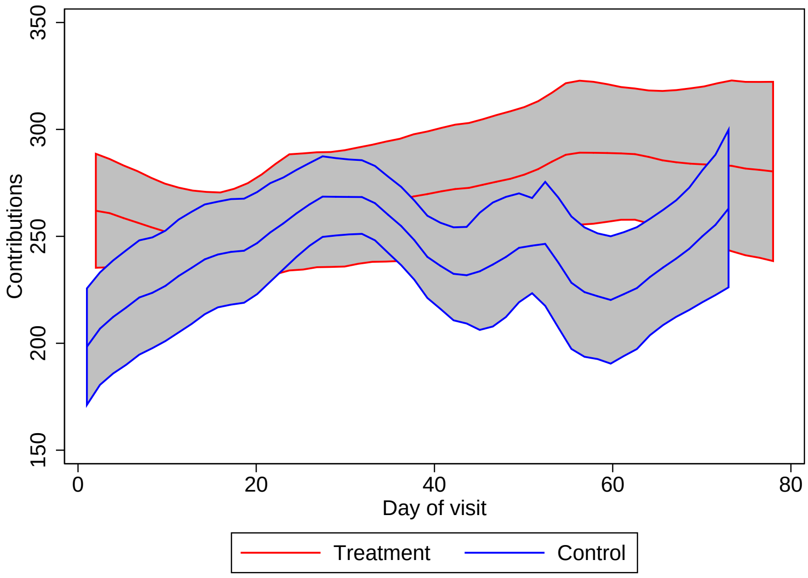
Analysis uses nearest neighbor propensity score matching, with 2 neighbors, with replacement. ATE of treatment on contributions. Significantly different from zero at * 0.1; ** 0.05; *** 0.01. Abadie-Imbens Robust Standard Errors in parentheses. Values of propensity score outside common support range are dropped. Exact matching excludes sectors with only 0 or 1 village in either treatment or control groups.

M Time, Field Team Experience, and Contributions

One possible concern is that the field team conducting the experiment gained experience over time, and this altered contributions among later participants relative to earlier participants. This could be a threat to identification of the treatment effects if there were a positive correlation between experience and contributions. Note however that, a priori, there does not seem to be any reason why the effect of experience would be positive or negative. In order to assess the potential for this concern to bias results, I examine the relationship of contributions over time (date of visit), by treatment and control groups separately. Date of visit ranges from 1 to 78.

Figure M1 presents Epanechnikov kernal-weighted local polynomial smoothing plots regarding the relationship between date of visit and contributions. In fact in the treatment group (villages with previous participating neighbors within 1.75 km) there is only a marginal increase in contributions over time, not statistically significant. This assuages concerns that the treatment effect is driven by increases in contributions due to field team experience, rather than due to communication with past participants. With the control group (no past participating neighbors) there do appear some patterns of increasing contributions over time, though this is non monotonic. This alludes to the possibility that some villages in the control group may nonetheless have had contact with previous participants.

Figure M1: Contributions over time



Epanechnikov kernal-weighted local polynomial smoothing plot showing relationship between date of visit and contributions.

Although visit date and treatment status are not significantly associated, they are positively correlated (F-test p-value 0.125), as one would expect given that villages visited later are more likely to have previous participating neighbors. One concern is that some of the observed treatment effect on contributions is being driven by a time trend, itself weakly correlated with treatment. To investigate this concern, first Table M1 presents the matching analysis with the inclusion of date of visit as a match variable. From this table one can see that the estimated average treatment effects are similar to the estimates in Table 2 in the main paper.

One can also add date of visit as a further explanatory variable in the OLS regression. Table M2 presents the analogous regression to Table 3 in the main paper, but with the addition of the date of visit. Date of visit is significant at the 10% level in column (1), but is no longer significant when further controls are added. and does not substantively alter the coefficient on the treatment dummy comparing with the estimates of Table 3 in the main paper.

Table M1: ATE of Presence of Past Participating Neighbors Within 1.75 km

	(1) Standard Matching	(2) Exact Matching
Contribution	25.418*	32.761***
	(13.997)	(10.836)
Observations	117	103

Analysis uses nearest neighbor propensity score matching, with 2 neighbors, with replacement. Significantly different from zero at * 0.1; ** 0.05; *** 0.01. Abadie-Imbens Robust Standard Errors in parentheses. Values of propensity score outside common support range are dropped. Exact matching excludes sectors with only 0 or 1 village in either treatment or control groups. Same matching strategy as Table 2, but adds the additional input variable of day of visit.

Table M2: Effect of Presence of Past Participating Neighbors Within 1.75 km

	(1)	(2)	(3)
Treatment Status	27.961***	25.120**	23.977**
	(10.323)	(10.289)	(11.742)
Date of Visit	0.399*	0.511	0.862
	(0.230)	(0.333)	(0.737)
Distance to base (km)		−0.396	−1.465
		(0.842)	(3.201)
Distance to paved road (km)		2.951	0.753
		(1.907)	(4.459)
Village Size (# HHs)		−0.011	0.077
		(0.117)	(0.138)
# Villages ≤ 1.75 km		1.016	3.634
		(4.753)	(5.626)
Years of Education		2.276	1.073
		(4.987)	(6.613)
Female		93.457***	82.731*
		(36.240)	(42.794)
Age		16.436	18.084
		(25.642)	(25.809)
Age^2		−0.218	−0.250
		(0.355)	(0.359)
Controls		✓	✓
Sector Fixed Effects			✓
R^2	0.07	0.22	0.30
Observations	147	147	147

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at * 0.1; ** 0.05; *** 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 1 in main paper.

N Experimental Protocol

In total 150 villages were randomly selected to participate in the broader health evaluation, and hence in the public goods games. Regarding individual selection, 12 individuals were chosen randomly from the baseline survey list, which included one person from every household in a selected village with at least one child under the age of 5, and were given a ticket to participate in the games the following working day. The team typically visited 2 or 3 villages per day. The average visit time of 3 hours included time spent tracking down the selected participants. At the time of the games, the 12 individuals were checked-in by the survey team and completed a brief questionnaire.[5]

The format of the games was standard, with a multiplier of three used to multiply the total contributions to the group. The endowment was 400 RWF, split into 4 coins of 100 RWF each. There were two paid rounds, with one round of practice, using paper tickets, conducted before the paid rounds. During the paid rounds, individuals were asked to go to a completely private area, one by one, and decide how much to contribute to the group fund. To make their contribution, each individual was given a small green coin purse, with which they could deposit their preferred contribution. The money they preferred to keep for themselves was simply kept on their person. This had both the advantage of being a natural place to store money, as well as making it salient that this money was theirs to keep. Each individual amount was recorded, using anonymous ID numbers located inside the contribution purses, to prevent identification of individuals by the survey team. Next, all the purses were emptied publicly one by one, and in a transparent manner the coins were counted, tripled, and divided equally among all 12 participants. Participants then received this income, and put it with the rest of their personal money.

The second round consisted of one of four different versions of the game, detailed in Section H and in the protocol below. Subjects were aware that there would be a second round, but were not given any information about the specific variation that would be used, ensuring comparability across villages.

Additional data notes:

- In two villages it was not possible to find 12 participants. These are excluded from the data.
- One village is missing from the final data.
- 14 individuals are missing from the final data. Their respective village averages have been calculated excluding their data.
- One village is missing data for number of households: this value has been imputed using the observed relationship between number of households and number of households interviewed as part of the community health project's baseline survey.

N.1 Participant Instructions

Note: Instructions were delivered in Kinyarwanda. Further, the final protocol was located on devices which have since been discarded. This protocol is based on the most recent recovered draft protocol, and has been updated accordingly. As such there may have been slight deviations from the wording found here.

[5] Eight individuals were randomly selected for a wait list, in case individuals did not show up at the specified time.

Rusizi Evaluation of Community Hygiene Clubs (REACH)
Participant's Script
Voluntary Contribution Exercises

[These instructions should be read out to the participants before commencing of the activity. Each group will only complete one of the four possible versions of Round 2]

Enumerator instructions: Ensure that the area is clear of spectators, and that each participant has been given the correctly numbered green purse.

Activity Description

Thank you for agreeing to participate in this activity. In this activity, you will be making decisions that can earn you money. The money you earn depends on your choices, as well as the choices of others. We will play the activity two times. Now we will explain the rules for the first time. The second time may be slightly different, and it will be explained when we get there.

Don't worry, we will answer any questions and conduct a practice before we begin.

As you can see, there are twelve of you in this room, and each of you has been given a green coin purse. Everyone will equally be given four 100 RWF coins (a total of 400 RWF).

The objective of the activity is for you to use the money provided in one of two ways. One option is to contribute some, or all of it, towards a group fund, by putting it in the green coin purse. The other option is to keep some, or all of it, for yourself. Regardless of what you do, this decision is private, and no-one will know what decision you choose.

All money contributed to the group fund will be added together, tripled, and then evenly split amongst all 12 participants, independent of their contribution.

Any money that you keep for yourself is directly yours to keep. Regardless of how much money you decide to keep for yourself, you will still share equally in the group fund.

Here is how you will make your choice of how much money (of the 400 RWF) to contribute to the group fund or keep for yourself. One by one, you will be asked to come to a private area, where no one can see your decisions. In this area you will allocate the money as you like. The money you choose to contribute to the group fund, you will place in the green coin purse. The money you will keep your yourself, you will keep on your person.

After making your allocation, you will bring your green coin purse to the "banker". After all 12 of you have dropped off your green purses, the banker will count all of the coins in all of the green purses, multiple them by three, and put an equal number of coins in every coin purse, and we will return them to you.

Before we start we will show an example.

Enumerator instructions: Go through the three examples with four mock participants. Example 1: Everyone contributes 0 to group, and everyone receives 400. Example 2: Everyone contributes 400 to group, and everyone receives 1200. Example 3: Three participants contribute 400, one contributes 0. The first three participants receive 900, the fourth receives 1300.

Are there any questions?

Now we will perform one practice round. Instead of 4 coins, now you will receive 4 tickets. Remember this is just for practice and not for money. Remember to respect the privacy of everyone and ensure there is no talking during the decisions.

Conduct full practice round.

Are there any questions? Remember to respect the privacy of everyone and ensure there is no talking during the decisions.

ROUND 1: BASELINE:

Enumerator instructions: One by one ask the participants to go to the private area to make their decisions. Participants will drop off green coin purses with the banker.

After all purses have been dropped off, the banker will count all the coins in front of all participants. Then the banker will triple (adding two coins for every one) the amount and return an equal amount back to every coin purse. Next each participant will receive their coin purse back and can put their earned money aside.

Note: Only one of the following four versions will be played:

ROUND 2: BASELINE (AGAIN):

Now we will play the second and final round of the activity. The second round is identical to the previous one. So we will now proceed.

Enumerator instructions: One by one ask the participants to go to the private area to make their decisions. Participants will drop off green coin purses with the banker.

After all purses have been dropped off, the banker will count all the coins in front of all participants. Then the banker will triple (adding two coins for every one) the amount and return an equal amount back to every coin purse. Next each participant will receive their coin purse back and can put their earned money aside.

ROUND 2: PENALTY:

Now we will play the second and final round of the activity. The second round is similar to the previous one. The only difference is that now, after everyone has made their contribution decisions, you will receive the green coin purse of another randomly chosen person in this room. This also means that someone else will have your green coin purse.

Without knowing who this person is, you will be able to see how many coins they contributed to the group fund. If you want, you will have the option to give this person a penalty. If you give them a penalty, the banker will remove 100 RWF from their earnings. But, giving a penalty will also cost you 50 RWF, which the banker will take from your earnings.

Remember, that someone will also see how much you contributed, and can choose to give you a penalty.

Before we begin, each of you will be given a ticket marked with an 'X'. To give a penalty to the person whose purse you hold, all you have to do is put this ticket inside the purse. If you do not want to give a penalty, keep the ticket, and do nothing.

Are there any questions?

Enumerator instructions: One by one ask the participants to go to the private area to make their decisions. Participants will drop off green coin purses with the banker.

After all purses have been dropped off, the banker will distribute these coin purses according to a pre-randomized order. Each participant will examine how many coins are inside the purse, and decide whether or not to give a penalty, by placing the ticket inside.

Next the coin purses are returned to the banker who will count the coins in front of all participants. Then the banker will triple (adding two coins for every one) the amount and return an equal amount back to every coin purse. The banker will remove 50 RWF for those who gave a penalty, and 100 RWF for those who received a penalty. Next each participant will receive their coin purse back and can put their earned money aside.

ROUND 2: REWARD:

Now we will play the second and final round of the activity. The second round is similar to the previous one. The only difference is that now, after everyone has made their contribution decisions, you will receive the green coin purse of another randomly chosen person in this room. This also means that someone else will have your green coin purse.

Without knowing who this person is, you will be able to see how many coins they contributed to the group fund. If you want, you will have the option to give this person a reward. If you give them a reward, the banker will add 100 RWF from their earnings. But, giving a reward will cost you 50 RWF, which the banker will take from your earnings.

Remember, that someone will also see how much you contributed, and can choose to give you a reward.

Before we begin, each of you will be given a ticket marked with an 'X'. To give a reward to the person whose purse you hold, all you have to do is put this ticket inside the purse. If you do not want to give a reward, keep the ticket, and do nothing.

Are there any questions?

Enumerator instructions: One by one ask the participants to go to the private area to make their decisions. Participants will drop off green coin purses with the banker.

After all purses have been dropped off, the banker will distribute these coin purses according to a pre-randomized order. Each participant will examine how many coins are inside the purse, and decide whether or not to give a reward, by placing the ticket inside.

Next the coin purses are returned to the banker who will count the coins in front of all participants. Then the banker will triple (adding two coins for every one) the amount and return an equal amount back to every coin purse. The banker will remove 50 RWF for those who gave a reward, and add 100 RWF for those who received a reward. Next each participant will receive their coin purse back and can put their earned money aside.

ROUND 2: UNCERTAINTY

Now we will play the second and final round of the activity. The second round is similar to the previous one. The only difference is that now, each of you will be given an additional blue coin purse. You will make your contributions as before, but this time there are two group funds, the green purse fund and the blue purse fund.

You can decide to contribute however you like – but money in one of these funds will be multiplied by 5. In the other fund the money will not change at all. Importantly, which fund is multiplied by 5 will be determined at random, by a coin toss, independent of your decisions. Therefore, you will not know in advance which fund is multiplied by 5, and which fund will stay the same.

The rest of the activity will be similar. After making all of your decisions in private, the banker will flip a coin to determine which fund is multiplied by 5. After this, the banker will count all of the coins in the green and blue purses, separately. Then, the selected funds coins will be multiplied by 5. The banker will add all the coins together, and divide them equally among all participants, exactly as before.

Are there any questions?

Enumerator instructions: One by one ask the participants to go to the private area to make their decisions. Participants will drop off green and blue coin purses with the banker.

After all purses have been dropped off, the banker will flip a coin in front of everyone. The outcome of the flip will determine which fund is multiplied by 5.

Next the banker will count the coins for each fund in front of all participants. Then the banker will multiply by 5 (adding four coins for every one) for the fund that was selected by the coin toss. The other fund will remain with the same number of coins. Next the banker will total all of the coins and return an equal amount back to every coin purse. Next each participant will receive their coin purse back and can put their earned money aside.

References

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Chris Hansen, "Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain," *Econometrica*, 2012, *80* (6), 2369–2429.