

# Identifying communication spillovers in lab in the field experiments

**Alexander Coutts**

Nova School of Business and Economics and NOVAFRICA

ISSN 2183-0843

Working Paper No 1903

February 2019

Revised October 2019

## **NOVAFRICA** Working Paper Series

Any opinions expressed here are those of the author(s) and not those of NOVAFRICA. Research published in this series may include views on policy, but the center itself takes no institutional policy positions.

NOVAFRICA is a knowledge center created by the Nova School of Business and Economics of the Nova University of Lisbon. Its mission is to produce distinctive expertise on business and economic development in Africa. A particular focus is on Portuguese-speaking Africa, i.e., Angola, Cape Verde, Guinea-Bissau, Mozambique, and Sao Tome and Principe. The Center aims to produce knowledge and disseminate it through research projects, publications, policy advice, seminars, conferences and other events.

NOVAFRICA Working Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

# Identifying communication spillovers in lab in the field experiments

Alexander Coutts<sup>‡</sup>

Nova School of Business and Economics

October 2019

## Abstract

Use of lab in the field experiments has increased dramatically, given benefits of studying relevant populations and their preferences. In the field, researchers must relinquish the control a standard laboratory offers, raising the specter of communication from past to future participants. While researchers may take steps to avoid such spillovers, less is known about their mechanics in these settings. In public goods games in rural Rwanda, I provide some of the first estimates of such spillovers by matching villages on all pre-study observables, comparing villages with and without communication opportunities. Communication led to substantial increases in cooperation, driven by conditional cooperators. The results suggest that spillovers are a real concern, and may bias inference about preferences. I conclude with advice to manage and test for bias from spillovers, and recommend that researchers increase transparency by adopting standard reporting procedures.

*JEL classification:* C92, C93, D83, H41, O10, O12.

*Keywords:* Lab in the Field, Public Goods Games, Field Experiments, Development, Rwanda, Communication, Spillovers, Information Transmission, Social Learning.

---

\*Nova School of Business and Economics, Faculdade de Economia, Campus de Carcavelos, Rua da Holanda, n.1, 2775-405 Carcavelos, Portugal; alexander.coutts@novasbe.pt

<sup>‡</sup>Acknowledgements: I would like to thank Hunt Allcott and David Cesarini for providing constant guidance and expert advice for this paper. I am additionally grateful for helpful comments from Isaac Baley, Canh Thien Dang, Deborah Goldschmidt, Nicole Hildebrandt, Elliot Lipnowski, Molly Lipscomb, David Low, Joseph Mullins, Yaw Nyarko, Rossa O’Keeffe-O’Donovan, Giorgia Romagnoli, Andrew Schotter, Emilia Soldani, Tobias Salz, Pedro Vicente, Christopher Woolnough, and seminar participants at CSAE, New York University, and Nova SBE. I am indebted to Zachary Clemence and Kris Cox at Innovations for Poverty Action for their help with management and implementation, and to Lea Stuff for assistance running the experiments. All errors are my own.

# 1 Introduction

Laboratory experiments to measure preferences and behavior have become a standard component of the economist's toolkit, and now a significant number of lab experiments are conducted in the field: outside of a university classroom or computer lab, and with non-student populations, following the definition of Gneezy and Imas (2017). The utilization of such lab in the field experiments has increased dramatically in applied work in the social sciences. In development economics, researchers are putting increasing weight on the importance of understanding relationships between preferences and development outcomes. Such work is increasingly finding outlets in leading journals of economics and political science.<sup>1</sup>

The internal validity of such studies hinges on unbiased identification of preferences, which among other things, requires thoughtful experimental design. There is significant awareness of the effects of individual learning both within and across lab games. Further, contrasting with lab in the field, field experiments are often designed with care to either prevent or account for spillovers, particularly when theory dictates, as with information interventions.<sup>2</sup> For example, malaria prevention information could spill-over to a control group, leading to a downward bias of the treatment effect of information. In contrast, lab in the field experiments often involve treatments within or across sessions which are unlikely to be salient enough to generate those type of spillovers. Rather, as in the focus of this paper, there may be spillovers from participation itself, an issue that has received scant attention in the literature.

I refer to such spillovers as inter-session communication spillovers, or simply communication spillovers for brevity in this article. They occur whenever participants in an experiment share any aspects of their experience, advice, or interpretation of the experiment with anyone else who has not yet participated, but is in any way connected with someone who will participate (including being a future participant). Part of the reason these spillovers have not received much focus is that standard economic theory typically rules them out.

---

<sup>1</sup>Recent examples include Avdeenko and Gilligan (2015), Casaburi and Macchiavello (2018), Enos and Gidron (2018), Jakiela and Ozier (2015), and Kosfeld and Rustagi (2015). Use of lab in the field experiments is also on the rise in other fields such as sociology, see Baldassarri (2015).

<sup>2</sup>I am describing field experiments in the sense defined by Duflo (2009), which List (2007) refer to as natural field experiments. Lab in the field experiments are alternatively referred to as artefactual field experiments. For further discussion of spillovers in field experiments see Banerjee et al. (2017) for further discussion, with examples in: Duflo and Saez (2002), Dupas (2014), and Miguel and Kremer (2004) among others.

A well-informed subject in an experiment should not be affected by advice or framing, as these behavioral phenomena do not alter the rational payoff-maximizing action. Yet the paradox is clear: experiments are interesting because they capture deviations from standard theory, thus spillovers which arise from behavioral impacts of communication are indeed a real threat. In experiments, behavior can change because of advice, altered beliefs about the behavior of others, or framing/experimenter demand effects.<sup>3</sup>

Yet, it is telling that in methodological accounts of lab in the field studies, communication spillovers are not mentioned as confounds that researchers need be aware of.<sup>4</sup> While lab in the field experiment methodology is less focused on spillovers than its field counterpart, often, though not always, researchers are cautious and adopt rules of thumb, such as trying to develop strategies to prevent contact between past and future lab participants.<sup>5</sup> At the same time, it is unclear the extent to which authors have in fact adopted these strategies: 3 of the 5 publications mentioned in the first paragraph do not explicitly address the potential issue of spillovers.<sup>6</sup> One of the most cited papers in economics using the lab in the field methodology, Henrich et al. (2001), was not immune to these concerns. In their phase II protocol, they raise concerns about contagion from communication spillovers

---

<sup>3</sup>A past participant may give advice to a future participant about what to do. They may reveal the outcome, thereby influencing expectations. Or they may give an interpretation, “I think they are trying to measure how cooperative we are”, which may create framing and experimenter demand effects. Schotter (2003) shows how advice can change behavior in experiments. Chaudhuri et al. (2006) shows how beliefs about others’ behavior can alter own behavior. Framing has been shown to affect behavior in lab experiments (Dufwenberg et al., 2011), as have experimenter demand effects, see Zizzo (2010).

<sup>4</sup>Neither Gneezy and Imas (2017) nor Harrison and List (2004) discuss spillovers as a potential threat to the validity of lab in the field studies. Viceisza (2012) does not discuss spillovers directly, but while discussing sampling recommends the scattering of households, subject to “reasonable limits”, which are not further defined. I am not aware of a methodological paper which includes the threat of spillovers across groups for lab in the field experiments. Within individuals, one may find many discussions of learning from repeated play in various games, a good starting point is Kagel and Roth’s *Handbook of Experimental Economics*. See Bednar et al. (2012) for learning across games.

<sup>5</sup>It is hard to know the true extent of the care taken to avoid spillovers in lab in the field experiments. In many cases it may not be obvious when spillovers may have posed issues. If they have posed issues, the compromised experiments may not end up being published. During presentations of this talk I have had multiple researchers describe experiences which involved communication spillovers requiring that experiments be re-done or thrown-out.

<sup>6</sup>Kosfeld and Rustagi (2015) conduct their lab in the field games with 56 different forest user groups, in five villages in Ethiopia. No mention is made of timing or location within these villages. Casaburi and Macchiavello (2018) conduct lab in the field experiments with farmers from a cooperative with 2,000 members, but do not mention the effect of timing or location of the experiments. Enos and Gidron (2018) conduct their experiment with randomly approached participants in cities in Israel, while not mentioning spillovers, their approach likely precludes such spillovers. Avdeenko and Gilligan (2015) mention geographical isolation to rule out spillovers. Jakiela and Ozier (2015) ensures communities are 5 km apart, to avoid spillovers.

in their behavioral experiments in 15 “small-scale societies”, discuss mitigation strategies, but conclude that the spillovers are impossible to avoid altogether.<sup>7</sup> Even if we assume adequate care is typically taken, the fact remains that we lack even the evidence to comment on the magnitude of the problem we are taking care to avoid. The bottom line is that failure to account for spillovers from such communication in experiments can bias the identification of preferences, which are often the main focus.

To a certain extent, the fact that individuals may communicate socially and learn from one another has always been possible in standard lab experiments. A typical participant in an economics lab experiment is often a student, and may have friends or colleagues who previously participated in a particular experiment. Thus, it is conceivable that these students may discuss outcomes or strategies of a particular experiment with friends who are about to participate. However, many experimental labs at universities may be less susceptible to this issue than corresponding field settings, such as rural Rwanda, the setting for this paper. For example participation in a university lab experiment may be less noteworthy, and involve lower stakes. Table 10 in the Appendix presents a quick summary showing that communication may be a bigger issue in lab in the field as well as framed field experiments.<sup>8</sup>

I study one such large implementation of a lab in the field experiment in Rwanda involving the participation of 150 different rural villages which was implemented over a period of three months. Theoretically, communication was not expected to alter behavior, although elements of the geographical study area, such as its relatively small size, suggested such communication may have been possible.<sup>9</sup> Observations from the field suggest anecdotal ev-

---

<sup>7</sup>Their protocol can be found here: [http://www.ensminger.caltech.edu/documents/587/3\\_-\\_Game-Procedures-and-Protocols-DG-UG-TPP.pdf](http://www.ensminger.caltech.edu/documents/587/3_-_Game-Procedures-and-Protocols-DG-UG-TPP.pdf). From their protocol: “Collusion/Contagion: Several members of our team have experienced some serious collusion among players in large villages where they have run significant numbers of experiments.”, and later, “You should be aware of this potential problem and be on the look out for it if you are returning to a village where you have played games before, or if you play a series of games this time in the same village.” One may be concerned if prevalence of spillovers was correlated with group characteristics such as “market-integration”.

<sup>8</sup>Many labs have multiple experiments occurring in a week, and often subjects are not aware which experiment they will be participating in. Additionally, economics experiments are not typically considered highly noteworthy events for a student. Comparatively, for participants in locations such as rural Rwanda, these events are perceived as being more out of the ordinary. Beyond this, financial stakes can be substantially greater in lower income countries; in the current experiment average earnings were greater than a typical full day of earnings.

<sup>9</sup>While this study should not be interpreted as representative of the average magnitude of spillovers in all lab in the field experiments, it should be considered important evidence on studies for which a reasonable threat of communication was present. While some elements of this study (small size of the region) suggested greater likelihood of communication, others (constantly changing location) suggested a reduced likelihood.

idence that communication had occurred. In one instance, the survey team visited a village that appeared similar to others in the region. Standard protocol was followed, however in this specific village, all (12) participants contributed the maximum possible amount in the public goods game. Because this was so exceptional, the team stayed behind to ask the villagers what led to such high levels of cooperation. A woman explained that she was friends with some of the women in a neighboring village, and one of her friends had participated in the same game only two days prior. Her friend told her that she should contribute the maximum amount, and she had shared this information amongst these villagers before the team had arrived.

To uncover unbiased estimates of communication spillovers, I utilize propensity score matching, which is well suited for this context. The reason is that in planning the order of visits of the 150 villages, the logistical planner for this study only had access to a pre-determined set of few village characteristics, and was personally unfamiliar with the villages. Thus, his ordering could only be conditioned on a small and finite set of observables, and not on private information. The matching strategy splits villages according to whether or not they had opportunities to communicate with past participants, using timing and GPS data, and then matches them on this key set of observables. The result is that one can compare villages which appear ex-ante identical to the planner, but for idiosyncratic reasons some were treated, i.e. they had neighbors which previously participated, and others were not. Having the full set of conditioning variables automatically fulfills the key selection on observables assumption underlying matching techniques, that treatment is independent of outcomes, conditional on observables. This in fact takes all the guesswork out of which variables need be included in the propensity score.

Implementing this matching strategy I indeed find large effects of communication. I find that being located near past participating villages increased contribution rates in the entire sample by 11-14% depending on the matching estimators used. I estimate these effects for different distance cut-offs, showing that spillovers cease to be significant after 2.5km. Beyond this, I consider two important robustness checks. First, in placebo tests, I consider how exceptional it would be to observe estimates of this magnitude for different counterfactual orderings of villages. The estimates of this paper exceed 99.9% of 10,000 simulated random orderings, and additionally always exceed those based on hypothetical deterministic orderings based on available observables such as the village's distance to the study base. Second, I use an instrumental variables strategy which also uses these deterministic orderings to generate a number of hypothetical past participating neighbors, which is then

used as an instrument for whether a village actually had past participating neighbors. The instrumental variables results suggest strong positive effects of communication.

While the evidence strongly suggests that communication had large impacts on cooperation, it is more difficult to identify the precise mechanism. A key explanation involves communication shifting beliefs about levels of cooperation upward, which would increase the contributions of so-called conditional cooperators: those who increase their contributions when they expect others to do the same. For example, Chaudhuri et al. (2006) find that providing future participants with public advice from previous participants in a public goods lab experiment increases average contributions. A similar mechanism appears to be at work in the field data: there is evidence that the treatment effect is driven entirely by those who are identified as conditional cooperators in the sample.

Evidence on unstructured social learning in lab in the field games is rare.<sup>10</sup> To my knowledge only Cardenas and Carpenter (2005) and Bernal et al. (2016) provide related evidence, from common pool resource lab in the field games in Colombia, where they conduct follow-up games with the same communities some months later. A critical difference is that these followup games included past participants, making up 30% of participants in Cardenas and Carpenter (2005) and 86% of participants in Bernal et al. (2016). While their focus is different from this paper, on the dynamics of learning within groups, the results are consistent: villages became more cooperative in the second visit. Importantly, Cardenas and Carpenter (2005) find that the results are not simply driven by past participants - novice participants are more likely to contribute more in the second visit compared with novice participants in the first visit.<sup>11</sup>

Thus, while researchers conducting lab in the field experiments may be concerned with communication across participants, the evidence is limited on the extent to which such communication would change behavior. This study presents the first evidence in a lab in the field context on how unstructured communication can alter behavior, and impor-

---

<sup>10</sup>On the other hand, in the lab there are studies of pre-play communication changing behavior, e.g. Isaac and Walker (1988a), see Sally (1995) for a broader overview, with more recent work by Bochet and Putterman (2009) and Brosig et al. (2003) for public goods games. It is not obvious that impromptu opportunities for communication during day to day life would lead to the same outcome as mandated and structured communication in a lab context.

<sup>11</sup>Both papers highlight the potential pedagogical effect of the experimental games, which could therefore have positively affected cooperative norms. These may be particularly salient in their setting, as in both papers the groups were selected specifically because common pool resources were important within the communities. Importantly, the mechanisms outlined in this paper cannot be excluded as alternative explanations.

tantly, the mechanisms underlying this flow of information. I document clear evidence of spillovers, find that they appear driven by conditional cooperators, and conclude that they pose a greater problem when participants are located within 2.5km of one another and when past participation occurred 3 to 7 days prior. The finding is significant, since these considerations exist in a context where standard theory makes the prediction that social information should not change behavior in public goods games.<sup>12</sup>

Traditionally, less attention has been given to social learning or spillovers in lab in the field experiments when compared with broader field experiments. The results and discussion of this paper lead to a number of practical suggestions for researchers either designing or analyzing projects involving lab in the field data. The first reiterates that care needs to be taken in design, as spillovers may arise even when theoretically unanticipated. The second is that techniques such as ex-ante randomization within geographical regions can be applied to generate ex-post tests for whether spillovers have occurred. And finally, attention must be paid to logistical planning. Even if logistics cannot be altered, the implementation plan can be key to uncovering estimates of potential spillovers, following the methodology of this paper.

To summarize the remainder of this paper, the next section outlines details of the public goods games and the data. This is followed by the matching analysis that demonstrates the effects of communication spillovers on behavior. I next spend time examining the role of conditional cooperators, and discussing potential alternative mechanisms, followed by a concluding discussion.

## 2 Implementation and Theory of the Public Goods Games

The experimental games were conducted as a component of a broader impact evaluation of community health programs.<sup>13</sup> In parallel with baseline surveys for that evaluation, public goods experiments were conducted in 150 villages in the Rusizi district in Rwanda. These villages were chosen randomly from a total of 598 villages in the district. For the purposes of this paper, the evaluation of the community health programs is not relevant, as these programs were implemented after the completion of all the public goods experiments.

---

<sup>12</sup>Standard theory posits a role for learning, in which case the learning should be about the Nash Equilibrium strategy to contribute zero, the opposite effect to what is observed.

<sup>13</sup>The official project name is: Impact Evaluation of Community-Based Health Programs in Rwanda (CBEHPP); ClinicalTrials.gov Identifier: NCT01836731; PI: James Habyarimana, Georgetown University. See Sinharoy et al. (2017, 2016).

The experimental games were conducted over a three month period in 2013. All 150 villages that were part of the larger evaluation participated in these games. 12 individuals were randomly selected from the household survey list, and given a ticket to participate in the games the following working day. The team visited either 2 or 3 villages per day, with each visit taking approximately 3 hours, which including 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.<sup>14</sup> Local survey staff then explained the game in the local language of Kinyarwanda. Communication during the game was not permitted, except to ask questions of the staff. A significant amount of time was spent explaining the game, including the tradeoffs between private and public benefits, providing a demonstration, and conducting a full practice session. It was important that individual decisions were completely private and anonymous; at no time were individual contributions revealed, a fact emphasized to participants.

The experimental design followed a standard public goods game format. Individuals were given an endowment of 4 x 100 RWF coins.<sup>15</sup> They were given real money to ensure that the stakes were salient, and to minimize confusion. One by one participants were instructed to leave the room, go to a completely private area, and decide how much to contribute to the common fund by depositing this in a small change purse, henceforth referred to as the contribution purse. The remainder of their endowment was kept on their person. This had the added advantage of making it clear for individuals that the money they kept was theirs. Since all individuals carry money on their person, there was no reason for concern about having decisions accidentally revealed.

After allocating their money, the participant would then place the contribution purse in a designated location. After all 12 participants had made their decisions, each individual amount was recorded, using anonymous ID numbers located inside the contribution purse, to prevent identification of individuals by the survey team. After recording, 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. Subjects played two rounds of the public goods game with real stakes, receiving income directly after each round. The second

---

<sup>14</sup>Eight individuals were randomly selected for a wait list, in case individuals did not show up at the specified time.

<sup>15</sup>At the time of the study 400 RWF was approximately 0.60 USD. From the Integrated Household Living Conditions Survey 2010-2011, 400 RWF comprises of more than an average day's income for 45% of the district population. Earnings in the experiment were larger than 1800 RWF on average, which greatly exceeds a day's income for the majority of the population.

round consisted of one of four different versions of the game.<sup>16</sup> 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. For this reason only the first round is used for the primary analysis of the paper, while the second round will be considered when examining the role of conditional cooperators.

Individual payoffs  $v_i$  were thus:

$$\begin{aligned} v_i &= 400 - c_i + \frac{3}{12} \cdot \sum_{j=1}^{12} c_j \\ &= 400 - \frac{3}{4} c_i + \frac{1}{4} \cdot \sum_{j \neq i} c_j \end{aligned} \tag{1}$$

Notice that contributing  $c_i = 0$  is the unique optimal strategy absent social preferences, with the marginal per capita return (MPCR) to cooperation set at 0.25. A long history of public goods experiments have shown that individuals tend to contribute non-negligible positive amounts in these type of games, despite the dominant incentives to free-ride. Chaudhuri (2011) provides a detailed overview of cooperation in public goods experiments, and places particular emphasis on the presence of individuals that play as conditional cooperators (CCs). CCs are so named because they contribute positive amounts in public goods games conditional on their beliefs about what others will be contributing. In their survey of the literature they find that CCs make up the largest group within experiments, representing between 35% to 81% of subjects.

CCs are important to be aware of as their optimal contribution depends on their beliefs about what the contributions of others will be, which will be an important potential mechanism for the present paper. Since this paper examines the potential effects of communication across participants in the public goods games, it is important to briefly review the evidence of how such communication could alter contribution decisions.

Chaudhuri et al. (2006) showed that CCs presented with public advice upwardly adjust their beliefs about average contributions, and thus increase their own contributions. They find that a substantial fraction of past participants advise future participants to contribute the full amount. When this advice is common knowledge, it has a significant impact,

---

<sup>16</sup>The four versions included the baseline game repeated, one game with the ability to punish, one with the ability to reward, and a unique game meant to measure uncertainty in public goods investment.

increasing contributions in the first round of play by over 18%.

Isaac and Walker (1988a) find that pre-game communication leads to higher contributions in public goods games.<sup>17</sup> Importantly, when decisions are private, CCs are essential for pre-game communication to have any impact on behavior. Because such communication is cheap-talk, any promises or announcements are non-binding. Hence of all the classified types of participants in public goods games, see Chaudhuri (2011), only CCs would be responsive to such communication, if it can successfully shift beliefs.<sup>18</sup>

To summarize, CCs make up the largest proportion of subjects in lab experiments and there is evidence that they are likely to be affected by either cheap talk or advice. Hearing a promise from an individual about contributing a large amount may lead one to increase their expectation of what that individual will contribute; receiving advice to contribute a large amount, and knowing that others have heard the same advice may have a similar effect. In order to identify CCs, in this paper I will use data on contributions in the second, final round, which occurred after participants had already observed first round contributions. Identifying CCs will be important when it comes to understanding the mechanisms at play.

### 3 Overview and Empirical Strategy

#### 3.1 Overview of the Games

Table 1 presents summary statistics of village level variables which were collected prior to participation in the games. The sample of participants was randomly selected from the population in the larger survey which included one person from *every* household in a selected village with at least one child under the age of 5. Women were over-represented in the larger survey, and subsequently are 74% of the participants in the experimental games.<sup>19</sup> The average years of completed education is 4.5, which corresponds to partially completing primary school, while the average age was 35.

---

<sup>17</sup>Crawford (1998) additionally surveys a number of experiments with cheap talk, finding similar patterns of effects.

<sup>18</sup>In the summary of Chaudhuri (2011), CCs are the most common type, followed by free-riders, then unconditional cooperators, and finally others who cannot be classified.

<sup>19</sup>Women were over-represented in this larger survey because the respondent was not randomly selected from within the household. Of note is that the gender ratio in Rusizi district is highly skewed, so the over-representation is not as large as it appears. According to the 2013 Rusizi District Gender Statistics Report, the percentage of females aged 25-64 is 55.4%. This age group represents 89% of those participating in this study.

Additionally, subjects were asked how many of the other 11 participants in the current session they knew well, by first and last name. Without this name requirement, piloting indicated that nearly all individuals reported the maximum, 11. I refer to these contacts as “strong ties”. The average number of strong ties was approximately 2.5. The two community indices were standardized responses to a survey question on to what extent they agreed that community members (1) cooperate with each-other, and (2) were willing to exert effort in the community. The trust index is a standardized response to whether they believe we can trust people or not (in general). Of the 150 villages visited, in only two villages were we unable to find the full 12 participants. These villages have been dropped from the analysis. The remaining sample is of 147 villages, with one additional village being unable to be matched.

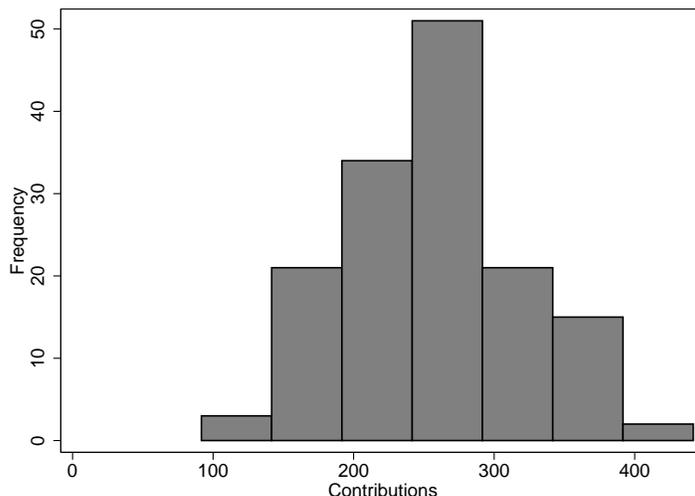
Table 1: Summary of Village Level Variables

	Mean	Std. Dev.	Min.	Max.
Average Contribution	254.92	65.89	91.7	400.0
Average Contribution (Round 2)	247.58	69.33	75.0	400.0
Proportion Female	0.74	0.14	0.4	1.0
Average Age	35.39	3.38	27.0	45.3
Average Years of Education	4.47	1.29	0.8	8.5
Average Number current Participants Known	2.56	1.09	0.7	6.8
Community Cooperation Index	0.81	0.13	0.3	1.0
Community Effort Index	0.79	0.16	0.3	1.0
General Trust Index	0.73	0.15	0.3	1.0
Distance to base (km)	13.14	7.98	1.0	38.8
Village Size (# HHs)	131.68	42.84	43.0	345.0
# Villages $\leq$ 1.75 km	1.86	1.33	0.0	5.0
Distance to paved road (km)	4.73	3.81	0.0	14.1
Observations	147			

Figure 1 presents the distribution of contributions in the public goods games. The possible levels ranged from 0 to 400 RWF, in 100 RWF increments. The average level contributed to the group fund is 255 RWF, which is about 64% of the socially optimal

level of contributing the maximum 400 RWF. As is typical in public goods experiments, the standard theoretical prediction of contributing zero is rejected.

Figure 1: Distribution of Contributions (RWF)



Contributions in the public goods experiments. Possible values ranged from 0 to 400 RWF, in 100 RWF increments.  $N = 147$ .

The 64% contribution rate is on the higher end of contributions in experimental public goods games. Typically, contribution rates range between 40-60%, though with a range of different MPCR. Another difference between these results and previous experiments is that the proportion of “free-riders” or those contributing nothing, is lower than commonly found.<sup>20</sup>

### 3.2 Strategy to Identify Communication Spillovers

The identification of the effects of communication on cooperation will involve a nearest-neighbor matching strategy, which pairs otherwise similar villages which only differed on whether there existed opportunities for communication to occur. Evaluating the validity of this strategy requires understanding the order in which villages were visited, since this ordering will determine which villages had opportunities to communicate with past partic-

<sup>20</sup>See Ledyard (1995) and Chaudhuri (2011). A MPCR of 0.25 is on the lower end, suggesting that the observed contribution rates are indeed quite high, see Isaac and Walker (1988b).

ipants. One ideal situation would be to have a completely random ordering. This would likely ensure that there is no differential selection of villages which had and did not have opportunities to communicate with past participants.

However in the current study the order was not explicitly randomized. Instead, a study planner observed some pre-study observable characteristics of the 150 villages that were in the study, and had to determine an ordering. Rather than assume (erroneously) that the planner chose randomly, I assume instead that after conditioning on all of the pre-study variables observed by the planner, the unexplained variation in order generates precisely the randomness needed for identification.

For this assumption to be credible, it would require that the planner did not have access to additional private information about the villages, beyond the pre-study observables. Based on interviews with the planner, this appears to be true. First, the planner was not a local of the study area, but was from a different district, and moved to Rusizi only for the duration of the project. Next, the planner was only personally familiar with two out of 150 villages in the study. In his role, he did not visit the villages, and hence did not have individual knowledge about them. His knowledge of Rusizi district was based on familiarity with the 18 political sectors which make up the district, their locations, and whether they are accessible by paved roads. Figure 5 in the Appendix presents a map of these sectors.<sup>21</sup>

Thus, the planner is unlikely to have had private information about the villages. How did the planner choose the order? As would be expected, the primary concern of the planner was logistical convenience: choosing to visit villages located nearby one another at similar dates. Two additional concerns featured in his decisions. The first was to ensure “difficult” villages were spread evenly throughout the study. Difficult villages were those that were located far from the study’s base location, and/or those that had large numbers of households, which required longer working hours.<sup>22</sup> The second was that on some days, the planner decided to alter the ordering due to heavy rainfall, as some regions were more difficult to access than others, due to a lack of paved roads.

In fact, these features of the planning process help to introduce exogenous variation

---

<sup>21</sup>The planner was interviewed in October 2019 by the author. He had been involved in supervising an earlier census of the project which tallied the number of households per village. As supervisor he did not personally visit the villages.

<sup>22</sup>All staff and enumerators were based in a central location. There were no overnight stays during the study, with the exception of 1 out of 18 political sectors. Removing this sector does not alter the results. The rationale behind spreading difficult villages evenly across the study was to minimize staff fatigue and ensure reasonable weekly working hours.

into the ordering. The first works to re-balance these characteristics among villages visited earlier and later, which may correlate with opportunities for potential communication. The second introduces a random shock, forcing the planner to alter the ordering. Overall, while most of the planner’s strategy was sensibly based on logistical convenience, there will be some variation in visit order, which cannot be explained in a purely deterministic way by all of the pre-study observable variables. In addition to these “shocks”, the fact that the planner lacked detailed information on the villages, but nonetheless had to make order decisions among villages which appeared similar from his perspective, means that necessarily some decisions would have to be arbitrary.

In the Appendix, Figure 6(a) presents the order of these visits on a map. The order is evidently not random, as neighboring villages are likely to be visited at similar points in time. What is important however, is that after conditioning on the observables available to the planner, the order exhibits random variation. To this end, Figure 6(b) in the Appendix shows visually the geographic distribution of villages which deviate from their predicted order, based on regression. As anticipated, the output does not present a clear pattern. The key strategy is thus to exploit the component of exogenous variation in the planner’s decision making, to find otherwise identical looking villages, based on all observables available to the planner, but by chance some had neighbors who previously participated in the public goods games, while others had no such neighbors. Those with previously participating neighbors thus had potential opportunities to communicate with past participants, while those without did not. It will also be important to control for total number of neighbors, this is discussed in further detail in the upcoming analysis.

For this estimation, it is most useful to consider an exercise where one frames this in terms of the treatment effects literature. In particular, certain villages are “treated” with exposure to previous participants of the public goods games, while others are not. This setting is particularly amenable to propensity score matching techniques, see Rosenbaum and Rubin (1983), in order to recover causal effects of communication on behavior. The reason is that the order of visit could only be conditioned on observables known to the planner, *before* the games were conducted. Thus one can make use of this full set of observables to generate propensity scores for the “treatment”: having neighbors who previously participated. By matching treatment and control villages with similar values of the propensity score, i.e. villages who are similar based on all observables available to the planner, one is able to recover causal estimates of treatment effects.

The key assumption in using matching techniques is the selection on observables as-

sumption, i.e. that after controlling for observables, treatment assignment is independent of the outcome of interest: contributions in the public goods game. In most empirical studies, this assumption is difficult to satisfy, since often the treatment of interest, e.g. job training, is self-selected into by individuals, and the econometrician may not observe all relevant variables that determine this selection. In the case of the present paper, the selection on observables assumption is satisfied automatically, as I have the small but complete set of possible conditioning variables.

### 3.3 Defining the Treatment

The treatment of interest for this study is whether a village had opportunities for communication with past participating village neighbors. The treatment is not the effect of communication directly, since this is not observed. Two villages are defined as neighbors if they are within 1.75 kilometers of one another.<sup>23</sup> This specific distance is chosen as it reasonably captures opportunities for cross-village communication, but additionally because it creates balance between villages with neighbors, and those without. As will be shown, this is important for power in the matching strategy.

However, since this research question and distance was *not* identified prior to the study, it is important to demonstrate that the results are consistent for other distances, and that 1.75 kilometers was not chosen after an ex-post comparison of different distances. To demonstrate this, Appendix Section 5.5 shows robustness checks which verify that the results are broadly consistent for other distances. Moreover, the main results of this paper are presented for the distance of 2 km in Online Appendix Section A, which was the initially chosen distance for this paper. The results remain significant and similar in magnitude.

There is a further dimension through which treatment definition may be interpreted differently, and that is the threshold for the number of neighboring villages. The primary concern is in identifying reasonable opportunities for communication. Using one neighbor is the most intuitive starting point. Moreover, it is important to note that choosing the threshold incorrectly should only work to bias the estimates downwards *against* finding effects. Choosing too low a threshold runs the risk of downward bias due to the inclusion of villages which didn't have opportunities to communicate in the treatment group. On the other hand, choosing too high a threshold runs the risk of downward bias due to control

---

<sup>23</sup>These are geodesic distances, i.e. as the crow flies. Given that the treatment distance cutoff is short enough to walk, this is likely to also be a good approximation of commuting distance. Obstacles such as mountains or rivers separating villages are rare, but could introduce additional noise in the measure.

villages having opportunities to communicate.<sup>24</sup>

According to the treatment classification of having at least one neighboring village that previously participated in the games, 74 out of 147 of villages are not-treated, while the remaining 73 are treated, resulting in a balanced 50% distribution, which will be important for the matching strategy. Figure 2 shows the number of villages in control and treatment groups, broken down into exactly how many past participating neighboring villages they had. Of note is that the vast majority (88%) of villages have 2 or fewer neighbours that participated. The maximum number of neighbours was 5, this occurred for only one village in the study. It is also of interest to note the geographic dispersion of treatment and control villages. Figure 3 presents a map showing every village in the study by treatment status. One can see that there are no obvious geographic differences by treatment status.

Figure 2: Defining the Treatment

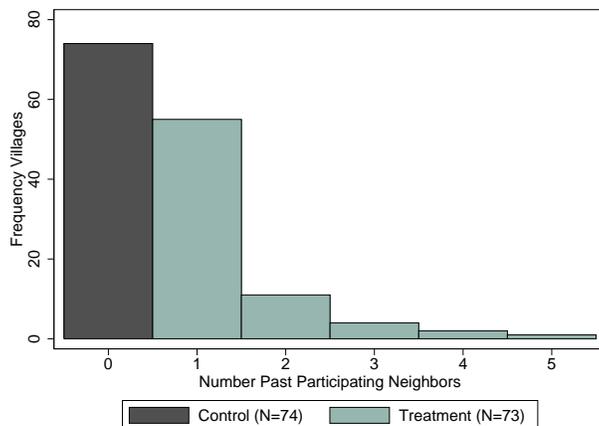
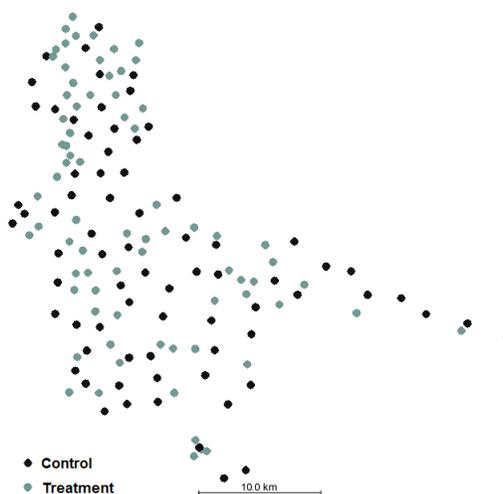


Figure shows the frequency of observing villages in the sample with given number of past participating villages within 1.75 kilometers as neighbors.

---

<sup>24</sup>There is also an element of timing. For example, one could further restrict the definition of treatment to only apply to villages that had neighboring participants in the last week or month. The relationship between information and time is theoretically ambiguous, as more time might allow information to be disseminated across villages or it could cause information to deteriorate. As the relationship may be non-monotonic, I remain agnostic by not conditioning on time. In the analysis, I examine heterogeneous treatment effects by days elapsed since neighbors participated, to empirically answer this question.

Figure 3: Map of Treatment Status of Villages



Each circle represents one village.

### 3.4 Determinants of visit order

Recall that the planner did not have familiarity with the villages, although he did know their geographic locations. Location will be used only indirectly in the matching strategy.<sup>25</sup>

The key variables available to the planner were the following:

1. Distance to village (from study base location)
2. Distance to paved road
3. Number of households in the village
4. Number of total villages in the study located within 1.75 km (village density)
5. Sector (political region; there are 18 sectors in the district).

---

<sup>25</sup>The reason for location to be used indirectly is that while ideally villages would be matched on distance from one another, this is not directly possible in the matching strategy. Moreover, distance to the base location and to a paved road will already convey much of this information, as will the addition of political sectors. Beyond this, adding latitude and longitude into matching may mis-characterize similarities (e.g. villages which share the same longitude but differ in their latitude or vice-versa). Given that the planner's knowledge was limited to the sector level, the analysis will give particular importance to additional ways of controlling for the sector.

Note that some of these variables contain information which may be used in ways that are difficult to capture in simple linear regressions, particularly regarding village locations. However, Table 11 in the Appendix examines such a regression of these variables on the order of visits. It can be seen that a simple regression with the four variables mentioned above entering linearly and sector fixed effects accounts for 96% of the variation in visit order. Thus one can conclude that a simple linear weighting of these variables is sufficient to capture nearly all of the decision making which uses these variables by the planner.

### 3.5 Summary Statistics on Treatment and Balance

Table 2 examines a logit regression village characteristics on the treatment indicator for potential communication, i.e. 1 if the village had at least one neighboring village within 1.75 km which previously participated in the public goods games, and 0 otherwise. This provides an overview of which variables are important for the determination of treatment.

From Table 2 it is possible to see that among the variables that the planner had access to, the only significant variable is the number of total neighboring villages within 1.75 km, independent of date of participation, i.e. the village density. This is not surprising, since villages with more neighbors that participated in the sample at any point in time are mechanically more likely to have neighbours at an earlier point in time.

Given this relationship, an issue could potentially arise if differences in contributions between treatment and control arise because of differences in village density. There are a number of ways I will address this possible issue with identification. First, one thing to note is that a small proportion, 16% of villages, had no neighbors within 1.75 km. Mechanically, these villages cannot be in the treatment group. As such, I remove them from the matching analysis.<sup>26</sup>

For the most part, matching will alleviate this issue as only similar villages across treated and control groups will be matched, and number of neighbors is a key variable in the matching estimation. Beyond this, in Appendix Section 5.5.3 I examine exact matching of villages by density, i.e. comparing only villages with the same number of neighbors within 1.75 km in the study, and show that the results continue to strongly hold. As an additional sanity check there is no statistically significant relationship between contributions in the

---

<sup>26</sup>If one includes these villages, 88% of them end up being removed anyways to due extremely low propensity score values that fall outside of the common support. As such, including them does not alter the main results of this paper.

public goods games and the number of sampled villages within 1.75 km.<sup>27</sup>

Table 2: Logit regression for treatment: Past participating neighbors within 1.75 km

Logit Regression	Treatment (1)
Distance to base (km)	0.041 (0.110)
Village Size (# HHs)	-0.005 (0.006)
# Villages $\leq$ 1.75 km	1.084*** (0.219)
Distance to paved road (km)	0.059 (0.153)
Observations	143
Sector Fixed Effects	✓

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors are reported in parenthesis.

Next Table 3 examines the balance across treated villages: those with at least one neighboring village which previously participated in the public goods games, and control villages. Only one statistically significant imbalance is found for variables which the planner had available, namely the number of total study villages located within the 1.75 km radius, echoing the earlier result. Distances to the study’s base location or to the nearest paved road are not significantly different across the groups, nor is the village size. Regarding the variables unavailable to the planner, the only statistically significant variable is in fact the main outcome of interest of this paper, village average contributions in the public goods game. In particular contributions are 30 RWF *greater* in treatment villages, i.e. villages which had neighbors that were past participants.<sup>28</sup>

Since treatment and control villages are similar on all observable characteristics related to demographics and preferences, one might in fact interpret the difference in contributions as the unbiased treatment effect, that is the true impact of opportunities for communication

<sup>27</sup>They are positively correlated. The p-value on an F-test for whether village density can explain contributions is 0.165. Adding the treatment dummy reduces this p-value to 0.868. Additionally village density is uncorrelated with contributions in the control group (F-test p-value 0.878). This relationship is investigated further using OLS regressions in Online Appendix Tables D2 and D3. Finally, an additional reason this is less of an issue than first appearances may suggest is that the variable is defined as neighboring villages *in the study*. Since only 150 villages out of 598 participated in the study, the variable itself is only correlated with the actual number of neighbors.

<sup>28</sup>Additional variables are explored in Online Appendix section B. These variables are related to health and household assets, coming from the baseline survey of the evaluation of community health programs. There are no statistical differences between treatment and control for these additional variables.

Table 3: Balance of treatment: Past participating neighbors within 1.75 km

	Treatment	Control	Difference
<i>Available to planner</i>			
Distance to base (km)	12.17	14.09	-1.92
Village Size (# HHs)	128.65	134.64	-5.98
# Villages $\leq 1.75$ km	2.48	1.24	1.24***
Distance to paved road (km)	4.96	4.50	0.45
<i>Unavailable to planner</i>			
Average Contribution	270.22	239.83	30.40***
Proportion Female	0.74	0.74	0.01
Average Age	35.39	35.38	0.00
Average Years of Education	4.53	4.40	0.13
Community Cooperation Index	0.81	0.82	-0.01
Community Effort Index	0.79	0.80	-0.01
Average Number Strong Ties	2.53	2.60	-0.07
General Trust Index	0.72	0.74	-0.02
Observations	73	74	147

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

on cooperation. However, it will be important to ensure that potential differences in selection of visit order, based on the variables observable by the planner, are controlled for in a more systematic way. The next section presents the matching strategy.

### 3.6 Propensity Score Matching

This setup is well suited to matching, due to the fact that all observables that the treatment could have been conditioned on are available in the data. While control and treatment villages may on average have differed on some characteristics, matching enables one to compare similar groups of villages who either received or did not receive the treatment respectively.

Here I follow the notation of Imbens and Rubin (2015), with some slight adaptations. Let  $C_i(1)$  denote the outcome of interest, village level contributions, if village  $i$  had at least one neighboring village within 1.75 km which previously participated in the games (treated,  $W_i = 1$ ), and  $C_i(0)$  be the contribution of a village with no previous neighbours

(untreated,  $W_i = 0$ ). In an ideal world, one could observe both outcomes (treated and untreated) for the same village, and hence could calculate the average treatment effect  $\tau$ . In the real world, the classic problem is that one cannot obtain an unbiased estimate of the treatment effect by naive comparison of the average outcomes of the two groups ( $\bar{\tau} = \bar{C}(1) - \bar{C}(0)$ ) because these groups may have different characteristics.

In practice, randomization can solve this problem, by creating comparable treatment and control groups. Here, randomization did not occur. Instead, following Rosenbaum and Rubin (1983) and a number of others, the strategy is to find a set of observable covariates  $X$ , which are known to be not affected by the treatment, such that:

$$W_i \perp C_i(1), C_i(0) | X_i. \quad (2)$$

This assumption is referred to as unconfoundedness or selection on observables. It means that the outcomes are uncorrelated with treatment, conditional on covariates  $X_i$ . In the current context this assumption is likely to be satisfied. The reason is that, unlike most observational studies, the treatment  $W_i$  (being exposed to villages who previously participated) could only have been conditioned on observables. This is because, as stated earlier, the planner determined the order of visits, in advance, with a limited number of pre-visit observables, and had no prior familiarity with the villages. In particular, it would be impossible for the planner to condition the treatment on features of data which had not yet been collected.

Denote the propensity score,  $e(x)$  by:

$$e(x) = Pr(W_i = 1 | X_i = x), \quad (3)$$

i.e. the probability that a village receives the treatment conditional on having characteristics  $X_i = x$ . This is also equivalent to the expectation of the treatment,  $\mathbb{E}[W_i = 1 | X_i = x]$ .

We can thus define the average treatment effect as:

$$\tau = \mathbb{E}[\mathbb{E}[C_i | W_i = 1, X_i] - \mathbb{E}[C_i | W_i = 0, X_i]] \quad (4)$$

As Imbens and Rubin (2015) note, in addition to unconfoundedness, a second key assumption is required for the analysis. This involves a requirement that there is overlap

in the distribution of covariates across treatment and control villages. Intuitively speaking, one needs to be able to find similar villages in control and treatment groups, in order to make valid comparisons.

Regarding this second assumption of overlap, as noted earlier, in Table 3, average characteristics on variables observed by the planner are reasonably balanced across treatment and control villages. As the next section will show in more detail, the assumption of overlap is broadly supported in the data.

### 3.7 Estimating the Propensity Score

The propensity score needs to be estimated from the covariates which may potentially have had an impact on which villages received the treatment (having neighbors that previously participated). In the case of this study, these variables can only come from the set of all observables available to the planner at the time the order of visits was determined. In determining the propensity score, I do not include sector dummies, as sectors will be conditioned on using an exact matching strategy, which is detailed in the next section.

Regarding the key variables available to the planner, outlined in section 3.4, one cannot assume that the planner used these variables in a linear way. Thus it is also important to take into account potential higher order interactions between these variables and the treatment. To determine the optimal specification, I follow the algorithm outlined in Imbens and Rubin (2015), which involves selection of these higher order terms based on their added value in terms of predicting treatment assignment. The algorithm involves step-wise regression estimation of the propensity score, to select only those covariates that add value in determining treatment status.<sup>29</sup>

In fact, the algorithm does not select any additional higher order terms. Thus the final terms selected for estimation of the propensity score are solely the four variables corresponding to distance from base, distance from paved road, number of households in village, and village density within 1.75 km.

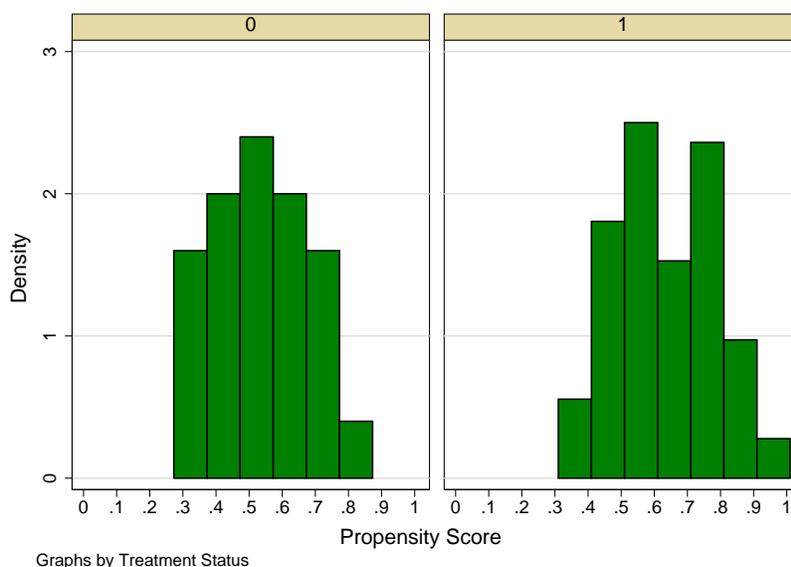
Figure 4 presents the distribution of the propensity score by treatment status. There is significant overlap over the sample, with the exception of values close to 1.<sup>30</sup> In the analysis I will impose restrictions that matching must occur in regions with common support.

---

<sup>29</sup>I follow Imbens and Rubin (2015), and set the threshold value for second order terms to be  $C_{qua} = 2.71$ .

<sup>30</sup>If one were to include villages with no neighbors, this would also present additional imbalances in the left tail.

Figure 4: Propensity Scores by Treatment Status



Distribution of propensity score.  $N = 123$ .

### 3.8 Results on Average Treatment Effect (ATE)

I focus on two main specifications for computing average treatment effects. Both include in the calculation of the propensity score the four key variables that the planner had available. The first specification does not utilize the 18 political sectors in the propensity score, in order to not dramatically increase the dimensionality. To account for the possibility that matching on sectors is important, the second specification forces exact matching on these political sectors. That is, I require that in addition to villages being similar across treatment and control according to available variables to the planner, I also require that these villages be located in the same political sector.<sup>31</sup>

Table 4 presents these two main empirical specifications. The first column presents standard matching estimates using two neighbors, with replacement. Alternative estimates for 1 or 3 neighbors are found in Appendix Table 12, with similar results. The second column presents exact matching on political sectors, again with matching on two neighbors.

<sup>31</sup>Given that the planner’s knowledge was at the sector level, exact matching will control for any private information the planner could have about these sectors.

The average treatment effect ranges from 28.4 to 36.9 RWF, each significant at the 1% level. These results are quite similar in magnitude, despite the specification differences, and correspond to a 11-14% increase in contributions over the entire sample.

Each approach has its advantages. Appendix Figure 10 presents the balance of covariates after matching. Covariates are balanced quite evenly across treatment and control after matching. In column 2 where exact matching on sectors is implemented, balance is only marginally affected. However, this comes at a cost of observations, as some sectors have few villages in either treatment or control groups. Regardless, it is re-assuring that both estimates are similar in magnitude.

In the Appendix, Section 5.5 examines these specifications for different treatment definitions that involve different distances. There Figures 7 and 8 present graphically ATE estimates for distances between 1 km and 3 km. The results are broadly consistent, with a slight pattern of shorter distances being associated with larger treatment effects, though there also arise issues of sample size, due to few treatment villages when distances are shorter, and few control villages when distances are longer. As noted, 1.75 km corresponds to the distance where treatment and control groups are most balanced. Importantly, after 2.5 km there do not appear to be any significant effects of spillovers.

One final concern could arise if the treatment effects are picking up positive effects of experience of the enumerators on contribution rates, or other interactions between the day of visit and village treatment status. With respect to the first concern, although it is not clear that experience would lead to greater (and not lower) amounts of contributions, I examine how contributions evolve over time, comparing the treatment and control groups. This is presented in Online Appendix Section I, showing that there is no significant increase in contributions in the treatment group. Online Appendix Section I also presents additional checks for including day of visit as both a control in the OLS specification, and as an input in the matching strategy. The variable is not significant in the OLS specification, and additionally does not substantively alter the matching estimates.

### 3.9 OLS

It is also useful to examine an OLS specification, particularly as the matching strategy reduces the number of observations quite substantially. Table 5 presents a simple OLS specification of the impact of having neighboring villages within 1.75 km who previously participated on contributions. It is possible to see that the impact is positive and significant,

Table 4: ATE of Presence of Past Participating Neighbors Within 1.75 km

	(1) Standard Matching	(2) Exact Matching
Contribution	28.368*** (10.677)	36.868*** (11.332)
Observations	117	101

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.

consistent with the matching estimates. The coefficient is approximately 30 RWF, a 12% increase from average contributions. Of note is that the estimate hardly varies at all when additional controls are added, including sector fixed effects. This is consistent with earlier evidence from Table 3 which demonstrated that the treatment was not significantly correlated with most observables. Beyond this, it is reassuring that the only variable significantly associated with treatment, the number of neighboring villages within 1.75 km, is not significant in the regression. Taken together, these results further suggest that treatment is indeed exogenous. Beyond this, the results obtained are quantitatively similar to those obtained using matching.

In addition to these results, Online Appendix Section C considers alternative dependent variables which capture opportunities for communication with past participants. These variables are (i) the order a village was visited, within its political sector; (ii) the number of villages in the sector which previously participated; and (iii) the distance of the closest previous participating village. All three are found to statistically significantly affect contributions, in line with the results found here. Additionally the Online Appendix Section D considers placebo style regressions for the number of total villages within the sector, as well as the density: number of villages within 1.75 km. Regarding this density, additional tests are made restricting the sample only to the control group. Reassuringly, in all cases, no statistically significant associations are found.

Table 5: Effect of Presence of Past Participating Neighbors Within 1.75 km

	(1)	(2)	(3)
Treatment Status	30.395*** (10.626)	30.876*** (10.604)	29.773** (12.032)
Distance to base (km)		-0.133 (0.850)	-1.312 (3.188)
Distance to paved road (km)		3.771** (1.786)	-0.307 (4.520)
Village Size (# HHs)		0.001 (0.115)	0.094 (0.142)
# Villages $\leq$ 1.75 km		-0.846 (4.770)	2.298 (5.676)
Years of Education		0.280 (5.502)	-3.146 (6.881)
Female		91.854** (40.884)	76.834* (45.056)
Age		10.552 (26.628)	15.083 (25.444)
Age <sup>2</sup>		-0.135 (0.368)	-0.206 (0.354)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.05	0.20	0.30
Observations	147	146	146

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 3.

### 3.9.1 Time and the Window of Communication

Of particular interest is the relationship between time elapsed since neighboring villages' participation and the contributions of treatment villages. Importantly, this relationship is ambiguous: more time between visits could allow information to reach more individuals, or it could lead the information to decay due to lost interest. The median of the average number of days elapsed since a treatment village's neighbors participated is 5 days, with a minimum of 1 and a maximum of 61. To explore this, I divide the treatment villages into three terciles: short (< 3 days), medium (3-7 days), and long (> 7 days). Table 6 presents the analog of the main OLS regression in Table 5, separating the treatment villages into these three groups. From Table 6, the strongest effects appear to be driven by the intermediate timing, of 3 to 7 days, showing a large effect that increases contributions

by 20%. Effects are also visible for shorter durations, but these are not always statistically significant. Finally, for villages which had neighbors that participated one week or more earlier, there is little evidence for strong effects, though the coefficient remains positive. While these results are highly suggestive that communication spillovers will have the largest impacts for intermediate dates (here 3-7 days), one must take into account the data limitations from such ex-post divisions of the treatment group.

Table 6: Effect of Presence of Past Participating Neighbors Within 1.75 km (By time elapsed)

	(1)	(2)	(3)
Treatment Status (short)	22.908* (13.269)	25.141* (14.241)	23.987 (15.437)
Treatment Status (medium)	50.445*** (17.873)	49.337*** (16.434)	50.827*** (19.140)
Treatment Status (long)	15.816 (15.929)	14.723 (16.093)	12.987 (20.439)
Distance to base (km)		-0.268 (0.880)	-0.940 (3.104)
Distance to paved road (km)		3.341* (1.735)	-0.445 (4.260)
Village Size (# HHs)		-0.012 (0.116)	0.097 (0.146)
# Villages $\leq$ 1.75 km		-1.129 (4.694)	2.771 (5.400)
Years of Education		0.592 (5.365)	-3.418 (6.750)
Female		97.238** (41.734)	78.662* (44.822)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.08	0.22	0.31
Observations	147	146	146

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 3. Short refers to the average date of neighbors' participation was  $<$  3 days, medium refers to 3 – 7 days, and long refers to  $>$  7 days.

### 3.10 Counterfactual Planner and Instrumental Variable Analysis

The results point to statistically significant sizeable effects of communication from past to future participants on contributions. Further, it appears that based on observables not available to the planner, villages that had neighbors which previously participated do not appear different from those that did not. Beyond this, matching and OLS strategies generate estimates which are similar in magnitude, suggesting that initial imbalances in the variables observed by the planner (namely the density of villages) do not appear to directly impact contributions. If any unobservables are correlated with these observables, this additionally assuages concerns that differences in unobservables drive the results.

Nonetheless, in this section I examine two additional robustness exercises, which will make use of hypothetical logistical paths for the planner. In the first exercise I compare the estimated treatment effects according to these hypothetical paths with the main results in the paper. This presents a statistical assessment of how unlikely it would be to observe the paper’s main results. The second exercise uses a subset of hypothetical paths to generate a set of instrumental variables for the treatment. I then instrument the observed treatment variable using these deterministically generated instruments, in a two-stage least squares (2SLS) analysis. The results of both these exercises support the main results, and the identification strategy of the paper.

Regarding the first exercise, to get a statistical sense of how unlikely it would be to observe the results of this paper, I generate two categories of different logistical paths for the planner, run the same statistical analyses, and compare these estimates with the main estimates of the paper. The first category involves simulating 10,000 random paths, defining the analogous treatment, and then examining how often an estimate is observed that exceeds in magnitude the estimates of this paper. I find that the main estimates in the paper exceed 99.9% of those estimates resulting from these simulated paths. This analysis provides a type of non-parametric test of significance. Further details, including plots of the simulated estimates can be found in Online Appendix E.<sup>32</sup> This exercise provides a measure of how “lucky” the planner would have to have been to by chance choose villages in such a way that treatment villages contributed the extent to which they did more than control villages. These results suggest that it is highly unlikely that the planner chose the

---

<sup>32</sup>By estimation strategy this is (1) 99.87% for standard matching, (2) 100% for exact matching and (3) 99.94% for OLS. Note that this is not identical to the permutation test where treatment is randomly re-assigned and exact tests of significance can be conducted. The reason is that treatment is not a monotonic function of visit order, and in particular depends on the spatial relationship between villages.

villages in the precise order which resulted in cooperative villages having past participating neighbors, and uncooperative villages having none. Particularly as the planner had no information about cooperativeness in the villages, nor other relevant private information.

The next category is to consider a number of “worst case scenarios”. I imagine a scenario where the logistical planner chooses the ordering in an extreme way - by simply ordering villages according to values of the available variables. For example, one scenario involves the planner choosing villages in order from the nearest to the study base to farthest. I do this for all variables (except one) available to the planner, ranking variables from either low to high and vice-versa (randomizing the order with ties).<sup>33</sup> The results of this exercise are shown in Table 7. As one can see, even in these extreme scenarios, the treatment effects are never significant and vary widely and inconsistently across the matching and OLS specifications. Taken together these results suggest that it is highly unlikely that the main results of this paper are driven by unobservables correlated with the order of visits.

Table 7: Counterfactual Tests

	(1) Matching	(2) Exact Matching	(3) OLS
Distance to base (km)	-12.067 (21.959)	10.924 (30.465)	-25.220 (15.870)
<i>Reverse order</i>	6.171 (15.979)	-8.690 (20.871)	-7.741 (14.860)
Distance to paved road (km)	-5.454 (13.408)	-15.880 (17.958)	-8.503 (12.831)
<i>Reverse order</i>	-17.304 (14.027)	19.884 (16.970)	-13.258 (14.159)
Village Size (# HHs)	-4.757 (20.977)	-17.919 (17.959)	-18.411 (14.615)
<i>Reverse order</i>	-18.468 (15.604)	-9.538 (15.521)	-6.399 (15.246)

Average Effect of Presence of Past Participating Neighbors, for counterfactual orders of village visits. Dependent variable is contributions. Order of visit is simulated for low to high values for odd rows, and the reverse (high to low values) for even rows. Analysis uses OLS regression and matching, following the empirical strategy in the main paper. Significantly different from 0 at \* 0.1; \*\* 0.05; \*\*\* 0.01. Abadie-Imbens or standard robust standard errors in parentheses respectively. Observations vary.

<sup>33</sup>The density of villages within 1.75 km is omitted from this analysis because of insufficient variation. In particular, ties between villages need to be broken randomly which results in variation in these estimates tied to this randomness.

The second exercise involves 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 generate a set of hypothetical routes that are selected based on deterministic functions of the observables to the planner. The next step is to define the number of hypothetically treated neighbors (within 1.75 km), which will be the instrument. Because the actual route was chosen largely based on logistical convenience (making use of these observables), it is reasonable to hypothesize that these hypothetical routes will lead to the number of treated neighbors being correlated with the treatment derived from the actual route.

Reminiscent of the first exercise, I calculate the six such routes which rank the observables of the planner from low to high, and vice-versa, as in Table 7. However in this case, I require that sectors are visited sequentially, i.e. that all villages in a sector are visited before moving to the next sector.<sup>34</sup> To choose among the resulting six 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. Further details are presented in Online Appendix E.<sup>35</sup>

Table 8 presents the results of the instrumental variables specification. The estimated coefficients are statistically significant at the 5% level, and significantly larger in magnitude than the matching or OLS estimates. In particular they suggest an effect of communication of about 49 RWF, or a 19% increase in contributions. One note of caution is that these estimates are less precisely estimated, as the standard errors are significantly larger the corresponding Table 5. Importantly, one cannot reject that the 2SLS estimate is equal to the corresponding OLS coefficient, for column (3) the corresponding p-value is 0.415.

---

<sup>34</sup>The reason is that without imposing such sequentiality, the resulting number of hypothetically treated neighbors is uncorrelated with the treatment, invalidating its use as an instrument.

<sup>35</sup>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. Full results are presented in Online Appendix E without this modification.

Table 8: IV-2SLS - Effect of Presence of Past Participating Neighbors Within 1.75 km

	(1)	(2)	(3)
Treatment Status	44.376*	48.024*	48.854**
	(24.175)	(25.246)	(21.377)
Distance to base (km)		-1.107	-2.170
		(0.961)	(2.997)
Distance to paved road (km)		4.101**	0.457
		(1.639)	(4.153)
Village Size (# HHs)		0.027	0.120
		(0.111)	(0.129)
# Villages $\leq$ 3 km		-4.471	-2.322
		(3.048)	(3.224)
Years of Education		0.137	-3.068
		(5.225)	(6.166)
Female		93.669**	87.118**
		(38.532)	(40.985)
Age		8.884	16.160
		(25.429)	(23.603)
Age <sup>2</sup>		-0.107	-0.216
		(0.352)	(0.328)
Controls		✓	✓
Sector Fixed Effects			✓
First stage F-Statistic	25.31	12.43	12.51
$R^2$	0.04	0.19	0.28
Observations	147	146	146

Analysis uses instrumental variable two-stage least squares (IV-2SLS) regression. Dependent variable is contributions. Treatment Status is instrumented with two instruments for the number of treated neighbors calculated from hypothetical routes, see Online Appendix E for details. Significantly different from 0 at \* 0.1; \*\* 0.05; \*\*\* 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table 3.

### 3.11 Mechanisms

I have shown in the previous analysis that individuals tend to contribute more in public goods games when they have neighbors who previously participated in the games. The next step is to more closely examine why this is the case. Most plausible, is that previous participants communicated with individuals who were yet to participate. In a standard theoretical model, pre-game communication should have no impact on behavior. Equilibrium contributions are expected to be zero, since subjects have full information about the

game and the determination of their payoffs. I now discuss four broad reasons why the theoretical predictions failed to materialize.

The first is a potential failure of the full information assumption if subjects do not understand the game. The second reason is that the theoretical game may be different from the actual game, as individuals may sanction outside of the scope of the game. The final two reasons relate to participants not behaving as standard rational agents. Specifically, the third is that playing the game may have altered cooperative norms, which were then passed on to future participants. Finally, the fourth and I argue most likely reason, is that individuals behave as conditional cooperators, and communication altered their beliefs about anticipated levels of cooperation. I now address these four reasons in detail.

If individuals do not completely understand the public goods game, this could lead to off path equilibrium behavior, i.e. contributing positive amounts. If pre-game communication improved individuals' understanding of the game, this should lead to decreasing contributions, not increasing as seen in the data. A possible explanation is that communication actually decreased understanding, leading to more confusion, and greater contributions. This seems highly implausible, and moreover, contributions in the second round of the game are statistically significantly (at the 10% level) lower, suggesting that learning may in fact lead to lower contributions.<sup>36</sup>

Next, the second reason is that there may be components of play that are not incorporated into the theory. A relevant example is social sanctions. If community members can punish one another outside of the framework of the game, it may be possible to sustain other equilibria. Collusion may lead to higher contributions, sustained by the threat of costly sanctions for deviators. Learning about the game in advance could also provide individuals with more time to devise a collusive strategy, which would explain the relationship between communication and increased contributions.

One convincing argument against collusive behavior is that there is no clear way to enforce such collusion. The implementation of the game placed a very strong and serious emphasis on privacy. Individuals were told that their decisions were private, and that they should not reveal their actions at any time during or after the game. While I do not observe

---

<sup>36</sup>At first glance it appears contradictory that communication can increase contributions, but individual learning can decrease them. Yet declining contributions over time are a robust empirical feature of public goods games, see Ledyard (1995). Additionally, Chaudhuri et al. (2006) find that in two of three treatments there is no significant correlation between one's own final contribution and one's advised contribution to others. They also find that advised contributions are always significantly greater than actual final contributions.

post-game interactions between players, anecdotal evidence suggests that individuals were not willing to reveal their decisions after the games. Collusion requires knowledge of individual decisions in order to credibly punish.<sup>37</sup> Since this information was kept private, it is unlikely that pre-game communication led to better opportunities to collude and hence higher levels of contributions. Additionally, as was mentioned, contributions were significantly lower in the second round. Playing a second round should generate additional opportunities to collude, implying contributions should increase or stay the same.

The third reason following Cardenas and Carpenter (2005) and Bernal et al. (2016), is that participating in the games may have positively altered community norms around cooperation. In the current context, this appears unlikely since past participants would need to convey these norms to future participants in other villages. This possibility can also be examined using the data, since the questionnaire included a question about whether people in the community were generally cooperative about issues that affect the community. Online Appendix F investigates whether having past participating neighbors within 1.75 km altered attitudes about cooperativeness in the community, and finds no effects. Hence it seems unlikely that cooperative norms were altered.<sup>38</sup>

Finally, the fourth reason is that the rational model may not be an accurate predictor of behavior, due to the presence of CCs. I focus on the two related behavioral phenomena introduced earlier: the effect of advice, and pre-play communication. Regarding the former, Chaudhuri et al. (2006) demonstrated in public goods experiments that when past participants gave public advice to future participants, this advice led to significantly higher contributions.<sup>39</sup> Regarding pre-play communication, Isaac and Walker (1988a) documented a significant role of such communication in increasing contribution rates in public goods games.

---

<sup>37</sup>One could make the argument that despite private decisions, individuals can make inferences based on the outcomes and priors about the cooperativeness of other participants. In this sense reputation and/or fear of reprisal can still operate. I thank an anonymous referee for making this point.

<sup>38</sup>The discussion below describes how it is likely that beliefs about contributions were shifted upwards when individuals had past participants as neighbors. This is not necessarily inconsistent with unchanged attitudes toward cooperativeness. Indeed, it would be surprising if simply the act of thinking about playing a game in the future, (until that point the game had not been introduced and even when introduced was couched in neutral language with no context), caused significant changes in how cooperative people believed their community was. One would expect that community-wide norms are much more stable than beliefs about behavior in an unfamiliar game.

<sup>39</sup>Note that there are differences between their controlled setting and this field experiment. First, their subjects are incentivized to provide advice, as they receive a portion of the future subjects' earnings. Second, in the common knowledge treatment where they find the largest effects, advice is read out loud in front of all participants.

Not surprisingly, I am unable to separately identify these two mechanisms. Presumably, when past and future participants communicated, some description of the experiment was given. But I do not know whether that description was bundled with advice. Nor is it reasonable to assume that every instance of communication followed the same structure. What I do in this section, is to identify which subjects are CCs, and examine whether villages with more CCs are more likely to increase their contributions as a result of having neighbors who previously participated.

As previously stated, all villages participated in a second round of the public goods games, that took the form of one of four variations, which I do not distinguish in this analysis.<sup>40</sup> I code individuals as being a CC if they contribute in the second round the nearest allowable amount to the modal contribution in the first round. If they contribute any other amount, I code them as not being a CC. Because the purses were emptied one by one and counted in front of all participants, the complete distribution of contributions was public knowledge. This procedure makes the mode particularly salient, though it will also be important to compare results when alternative definitions of CC using the median or mean are used. Aggregating this variable to the village level, the average proportion of CCs is 0.35, ranging all the way from 0 to 1 in the sample of villages.<sup>41</sup>

To examine the effects of being classified as a CC, the matching strategy is not tenable for splitting the sample into further subgroups, due to insufficient observations.<sup>42</sup> Instead, I examine heterogeneous effects using OLS regressions. Given the similarities between the matching and OLS results, and that only one covariate (village density) is imbalanced across treatment and control groups, such an OLS specification is likely credible. Table 9 presents the results interacting the treatment with the proportion of individuals classified

---

<sup>40</sup>One of the four variations was simply the baseline game repeated, which was intentionally over-sampled and repeated for 52% of all villages. Online Appendix Section G shows that the results are consistent selecting only the villages which repeated the baseline game.

<sup>41</sup>Hartig et al. (2015) is one of the few papers to study what conditional cooperators actually condition on. They find significant heterogeneity, but that subjects react positively to information on individual contributions, particularly when these contributions are greater and when there is less variation. This also hints at an important role for the mode of the distribution. Chaudhuri (2011) notes that experimental studies have found between 35% and 81% of subjects are CCs, but differences in classification procedures make it difficult to compare. An additional concern is that the proportion of CCs may itself be affected by communication. Reassuringly, the difference in the proportion of CCs is not statistically different across treatment and control villages (Ranksum p-value 0.2931). I thank an anonymous referee for pointing out this potential concern.

<sup>42</sup>Such an analysis is nonetheless conducted in Online Appendix H, however with sample size ranging from 19-57, there is not sufficient power to draw sharp conclusions, though the pattern of results is quite consistent.

as CCs in the village. As one can see, the effect of the treatment disappears, and the interaction becomes large and positively significant. The theoretical maximum treatment effect for village consisting of 100% CCs is 78 RWF, while the average effect given that villages have on average 35% CCs is approximately 21 RWF. Additionally, CCs contribute significantly greater amounts in general, approximately 88 RWF, or 35% more than their non-CC counterparts. This is intuitive, as typically after CCs, the largest proportion of classifiable subjects are free-riders, see Chaudhuri (2011). Table 14 in the Appendix examines different definitions for CCs involving medians or means, where similar effects are found for medians, while for the mean the magnitude is similar but the coefficient on the interaction is not significant.

The differential effect for CCs lends some support to the hypothesis that pre-game communication altered beliefs, leading individuals to increase contributions. However, I am unable to disentangle the relative importance of cheap-talk versus advice in altering beliefs. Overall, this result suggests that communication spillovers led to increases in contributions, and that these increases were disproportionately driven by CCs. This finding, combined with the previous discussion regarding alternative mechanisms, suggests that communication did not lead to a better understanding of the game, and likely did not create opportunities for collusion due to the private nature of decisions. Rather, it suggests that social learning altered subjects beliefs, as a result of pre-play communication, advice, or some combination of the two.<sup>43</sup>

Additional OLS specifications which split the sample into higher and lower proportions of CCs are presented in the Appendix. While heterogeneity on propensity to be a CC is central to this paper, it is also of interest to examine heterogeneity along other dimensions. Appendix Section 5.5.6 examines whether there are similar differences in treatment effects when median splitting villages by age, education, number of others known, and trust. However significant differences are only found for the specification for CCs.

Finally, one may wonder whether the outcome of the games for past participating neighbors matters for future participants. In theory the net impact of poor outcomes on beliefs is ambiguous. On one hand, poor outcomes may lead to pessimistic communication, lowering beliefs about overall cooperativeness. On the other hand, poor outcomes could

---

<sup>43</sup>Given the cultural context of these games, it is highly likely that if communication involved giving advice, that such advice became public and was widely shared amongst future participants. The woman described in the introductory anecdote explicitly stated that she had passed on this information to her co-participants. The common knowledge nature of this advice was shown to be a key component of increasing contributions in Chaudhuri et al. (2006).

motivate providing even more positive or urgent advice to future participants, e.g. “don’t make the same mistake we did”, which could increase expectations of cooperation among recipients of advice. This is difficult to answer with the data, as only 75 villages have past participating neighbors within 1.75 km, and there may be unobserved similarities for neighboring villages. In fact, comparing contributions in the treatment group based on whether the nearest past participating neighbor contributed above or below the median, the level of contributions is nearly identical: 270 and 269 RWF, respectively. Hence there is no evidence that the precise outcome of past games played a role in altering cooperation.

Table 9: The Role of Conditional Cooperators (CCs)

	(1)	(2)	(3)
Treatment Status	-3.897 (13.106)	-3.332 (13.120)	-10.246 (15.550)
Conditional Cooperator (CC)	106.185*** (30.387)	92.484*** (29.288)	88.166*** (25.901)
CC × Treatment	71.983** (33.518)	80.091** (32.043)	88.669*** (31.976)
Distance to base (km)		-0.410 (0.679)	-3.349 (2.558)
Distance to paved road (km)		1.075 (1.356)	-0.680 (3.862)
Village Size (# HHs)		-0.014 (0.101)	0.088 (0.114)
# Villages ≤ 1.75 km		-3.558 (3.355)	0.672 (4.052)
Years of Education		-3.066 (4.702)	-6.332 (5.075)
Female		62.474* (32.701)	33.779 (33.382)
Age		-2.557 (18.432)	4.352 (18.510)
Age <sup>2</sup>		0.036 (0.255)	-0.060 (0.259)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.43	0.50	0.58
Observations	147	146	146

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at \* 0.1; \*\* 0.05; \*\*\* 0.01. Robust standard errors in parentheses. Treatment defined as having past participating neighbors within 1.75 km. Controls includes all remaining variables found in Table 3.

## 4 Discussion

Lab in the field experiments have become increasingly important for research in economics and political science, as a means to study behavioral preferences and how they relate to broader economic outcomes. Yet field contexts may generate opportunities for communication between past and future participants to a greater extent than lab contexts. This can be problematic, when opportunities for communication are correlated with other unobserved characteristics of individuals or the environment, which themselves are related to outcomes under study. The results in this paper show convincing evidence that for public goods games in Rwanda, communication took place and changed behavior. Fortunately for identification, the opportunities for communication appear to be uncorrelated with other variables of interest.

Researchers are often aware of these problems, and may have experienced their effects firsthand. How should lab in the field experiments be conducted to minimize the effects of unintended communication on outcomes? The first relevant point involves the type of experiment and the context, and whether communication is likely to be an issue. This can be difficult to assess ex-ante. In the current paper, it was not anticipated that communication would take place. Moreover, since subjects were given full details about the rules of the public goods game, including a discussion about the tradeoff between private versus public payoffs, it was thought that even if there was communication, it would not change behavior. However there are some characteristics of Rusizi district of Rwanda which indicate communication may occur, namely that it is relatively small region, and has a relatively high population density ( $420/\text{km}^2$ ). Moreover, the median village has 2 other neighboring villages within 1.75 km, and this only refers to villages in the study. This makes it clear, that communication should not be unexpected. Thus researchers must assess whether communication is likely to be an issue in their context.

Next, researchers can take steps to control or mitigate the effects of communication. At the study design stage, this can involve creating suitable distance between sessions. If sessions occur in the same place, this becomes more difficult to control, and it may be necessary to avoid overlap in times, recruit from disparate populations, or create slightly different game versions which make it clear that advice from past games may not apply. The results contained in this paper suggest that there may be benefits to waiting at least 7 days between sessions.

Beyond this, a number of steps can be taken to test for and recover estimates of the

effects of communication, as in the case of this paper. In the case that sessions are in the same location, one can randomize the order, and see whether the order of the session has any impact on behavior. In the case that sessions are conducted in different locations, one can similarly randomize the order. Clearly, this may sometimes not be feasible due to logistics. In this event, one can select smaller zones of randomization: e.g. randomize the order within 10x10 km grids, or randomize the order within smaller political regions. In this manner one can test whether the order of visit within zones has any significant impact on outcomes.

Finally, if the design and order cannot be altered by the researcher, or if the study has already taken place, one can conduct an analysis following the techniques in this paper. The key part of this strategy is to pay attention to how the order of visits is determined, and identify all variables which have been used in its determination. In the case of this paper, the key to identification was leveraging the fact that logistical restraints required that villages participated at different dates, and because of natural variation in geography this led to variation in opportunities for communication between past and future participants. The matching strategy ensured balance on all covariates available to the planner at the planning stages, though the results suggest that imbalances do not pose a problem to identification in this context. I demonstrated that theoretically unanticipated communication can have large effects, representing increases of 11-14% in contributions to the public goods game.

These magnitudes should give any researcher pause in designing lab in the field experiments which may create opportunities for communication between past and future participants. Perhaps most importantly, is that researchers must adopt more transparent reporting practices regarding the details of their experiments, including the locations of sessions, the average time between sessions, and details about the local populations, following Table 10. If these details suggest cause for concern, then further analysis can be conducted by the authors using some of the suggestions in this paper. As current reporting norms appear to vary widely, this policy could have an important impact, through improving scholarly work, but additionally through raising awareness of these issues for future researchers.

Finally, as with spillovers in broader field experiments, context is important, and generalizability may depend on numerous factors. For instance, the effect of communication spillovers will likely depend on the nature of the game: games which present social versus individual tradeoffs may lead to more normative communication, whereas individual

decision making games may lead to more positive communication. Local norms may also interact in important ways with the content of such communication. Researchers need to be aware of the risks when designing lab in the field experiments, and aware of the means to avoid them. This paper presents some of the first evidence confirming that such communication spillovers can significantly bias the measurement of preferences in important ways, and illuminates some of the mechanisms at work, so researchers can take evidence based approaches to mitigating the problem.

## 5 Appendix

### 5.1 Threats of Communication Spillovers by Experiment Type

Threat	Description	Lab Exp.	Lab in the Field Exp.*	Framed Field Exp.†
Saliency	More salient/noteworthy activities are likely to be more discussed.	–	+	+
Short Distances Between Sessions	A fixed location or shorter distances provide more opportunities for communication.	+	+/-	+/-
Small Population	A smaller general population increases the likelihood of communication.	+/-	+/-	+/-
Greater Population Connections	A more connected population increases the likelihood of communication.	+/-	+/-	+
Stakes	Higher stakes are likely to increase the saliency of the activity, leading to more communication.	–	+	+
Number of threats beyond Lab Exp.:		0	+1-2	+2-3

Table 10: Table presents a list of potential threats which could be expected to increase inter-session communication spillovers in different categories of experiments. \*Lab in the field experiments are sometimes referred to as artefactual experiments. †Framed field experiments are described in Harrison and List (2004), as identical to artefactual field experiments but with field context in either the commodity, task, or information set that the subjects use. They often involve particular subject populations, who may be more likely to know each other, see for example Giné et al. (2010). This table does not include broader types of field experiments (e.g. natural field experiments), which don’t have straightforward definitions of “sessions”. Yet broader field experiments can also have spillovers from participation, in the sense that individuals exposed to some treatment may communicate their experience, advice, or interpretation of that treatment, which could alter even other treated individuals. For one example of this, see Dupas (2014) who finds that purchase of bed nets is higher among those treated with a subsidy when they were located near others who had been treated.

## 5.2 Political Sectors in Rusizi

Figure 5: Rusizi District Sectors



Map presents 18 political sectors of Rusizi district. The district is bordered by both DRC and Burundi. Nyamasheke is the bordering district within Rwanda. Map is adapted from Fourth Population and Housing Census, National Institute of Statistics of Rwanda.

### 5.3 Determinants of Visit Order

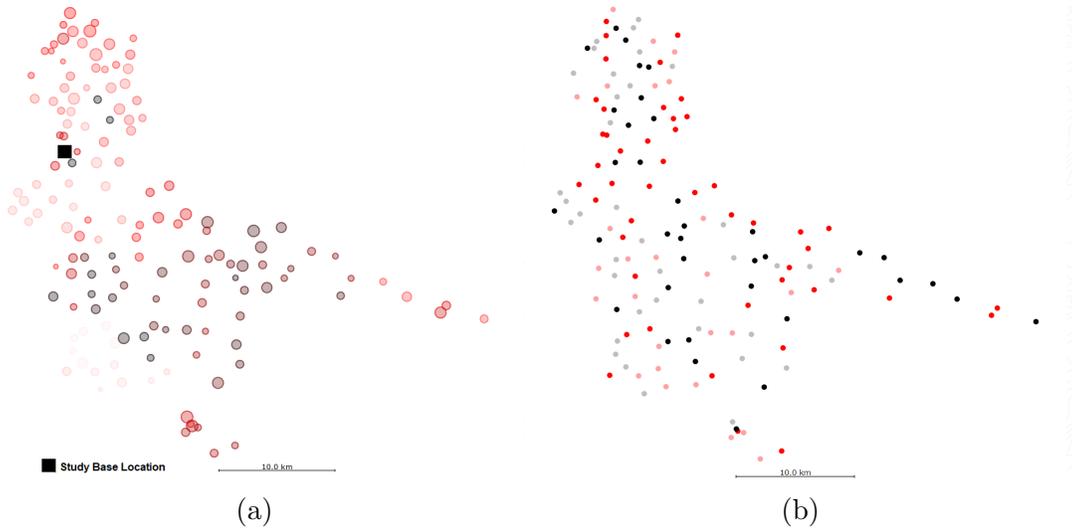
Table 11 presents an estimation of the order of visits, conditional on the variables observed by the planner (including the sectors). Importantly, this regression reveals that with an  $R^2$  of 0.96, these variables are sufficient to explain 96% of the variation in the order of visits. Next, Figure 6(b) presents the residuals from this regression visually. In particular, it shows deviations from the predicted ordering based on the variables observed by the planner. Villages visited 1 day earlier or more are shaded in black (this is 25% of the sample). Villages visited 1 day later or more are shaded in red (32% of the sample). And those that are visited within 1 day of their predicted value are shaded in light gray or red depending on whether they were visited slightly earlier or later respectively.

Table 11: Determinants of Visit Order

	(1)
Distance to base (km)	0.281* (0.155)
Distance to paved road (km)	0.123 (0.165)
Village Size (# HHs)	0.009 (0.006)
# Villages $\leq$ 1.75 km	-0.132 (0.199)
Sector Fixed Effects	✓
$R^2$	0.96
Observations	146

Analysis uses OLS regression. Significantly different from 0 at \* 0.1; \*\* 0.05; \*\*\* 0.01. Robust standard errors in parentheses. Dependent variable indicates order of visit and varies from 1 to 53. Villages visited on the same day receive the same value of this variable.

Figure 6: Village Visits



Each circle represents one village. (a) The shading constitutes the date of visit, with lighter circles representing earlier visits in the study, and darker later visits. Note the shading is not fine enough to display local variation in date of visit. The size of the circle corresponds to the average amount contributed by that particular village: larger circles correspond to greater average contributions. (b) Stronger shading corresponds to more unpredicted visit timing, defined by residuals of the regression in Table 11. In particular, black shading indicates villages visited 1 day or more earlier than predicted, while red shading indicating villages that were visited 1 day or more later than predicted. Light gray and red indicate those that were visited less than one day earlier or later, respectively.

#### 5.4 Different Numbers of Neighbors for Matching Estimation

Table 12: Average Effect of Presence of Past Participating Neighbors

	(1) Standard Matching	(2) Exact Matching
1 Neighbor		
Contribution	26.342** (11.041)	31.608*** (11.422)
3 Neighbors		
Contribution	30.814*** (10.298)	31.657*** (10.894)
Observations	117	101

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.

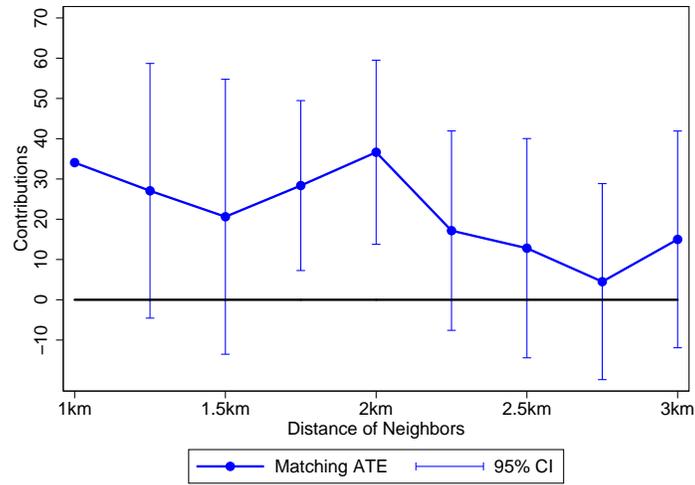
## 5.5 Estimates for Different Treatment Distances

### 5.5.1 Matching Estimates

Figure 7 additionally presents matching estimates for different distance cutoffs, using the specification with full controls analogous to column (1) in Table 4. The coefficient estimates range from 4.5 (not significant) to 36.6 (significant at 1%).

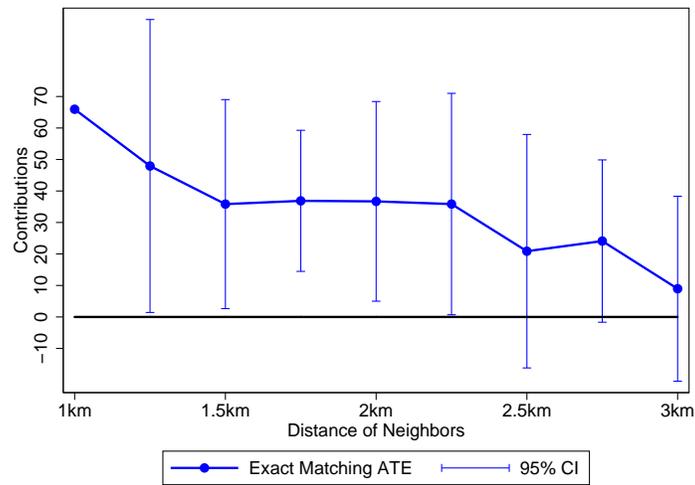
Figure 8 presents the analogous figure for exact matching on political sectors, the specification in Column (2) in Table 4. There one can see a pattern of diminishing effects as the distance cutoff is greater, where estimates range from 66.0 to 9.0 (neither significant). For the ranges 1.25km to 2.25km, estimates are significant at the 5% level. Note also that for these distances, treatment and control groups are also better balanced. Another point to note is that similarly decreasing patterns are observed for the OLS estimates in Figure 9, and one can see that OLS and matching estimates are similar for intermediate distances where standard errors from matching are moderate.

Figure 7: Matching estimates for different treatment cutoff distances



Each point corresponds to estimate of coefficient on treatment for specified distance cutoff for independent matching estimations, analogous to specification (1) in Table 4. Error bars suppressed for 1 km due to noise. Number of observations varies.

Figure 8: Exact matching on sector estimates for different treatment cutoff distances

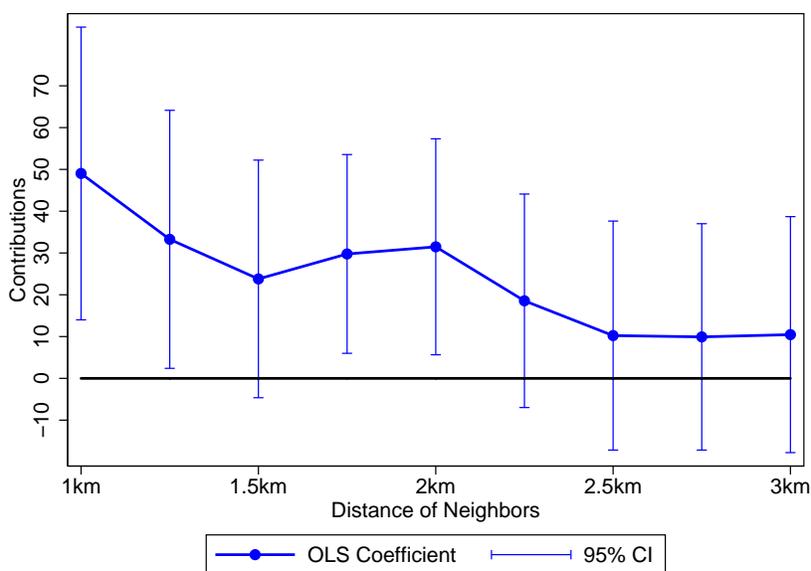


Each point corresponds to estimate of coefficient on treatment for specified distance cutoff for independent matching estimations, analogous to specification (2) in Table 4. Error bars suppressed for 1 km due to noise. Number of observations varies.

### 5.5.2 OLS Estimates

Figure 9 additionally presents OLS estimates for different distance cutoffs, using the specification with full controls analogous to column (3) in Table 5. These results show a pattern of diminishing effects as the distance cutoff is greater. The largest effects are found for villages with neighbors within 1km, increasing contributions by 49.0 significant at the 1% level, and the smallest at 2.75 km, increasing contributions by 9.9 RWF, not significant at conventional levels.

Figure 9: OLS estimates for different treatment cutoff distances



Each point corresponds to estimate of coefficient on treatment for specified distance cutoff for independent OLS regressions.  $N = 146$ .

### 5.5.3 Exact Matching on Village Density

In this section I mitigate concerns that treatment effects are in fact picking up differences in village densities, i.e. the number of neighboring villages in the study located within 1.75 km. I thus conduct an exact matching on the number of villages (in the study) that are within 1.75 km of a given village. By construction, the treatment variable, which is 1 whenever a village had neighbors located within 1.75 km that previously participated, is

highly correlated with the number of total neighbors within 1.75 km in the study. Further, the initial balance checks indeed revealed that village density was not balanced across treatment and control villages.

To account for the possibility that village density may be correlated with unobserved variables, and that the earlier propensity score matching may have been unable to adequately control for this, I conduct a matching strategy where I require that matched treatment and control villages *must* have exactly the same number of neighbors in the study. Village density has a minimum value of 0 and a maximum of 6, though values of 0 were excluded as noted in the primary analysis. Further, as there are no control villages for densities with 5 or 6 neighbors, these 5 villages are necessarily dropped from this analysis as well. Table 13 presents this matching analysis. As can be seen, the effects are significant, and consistent with the main results.

Table 13: Average Effect of Presence of Past Participating Neighbors

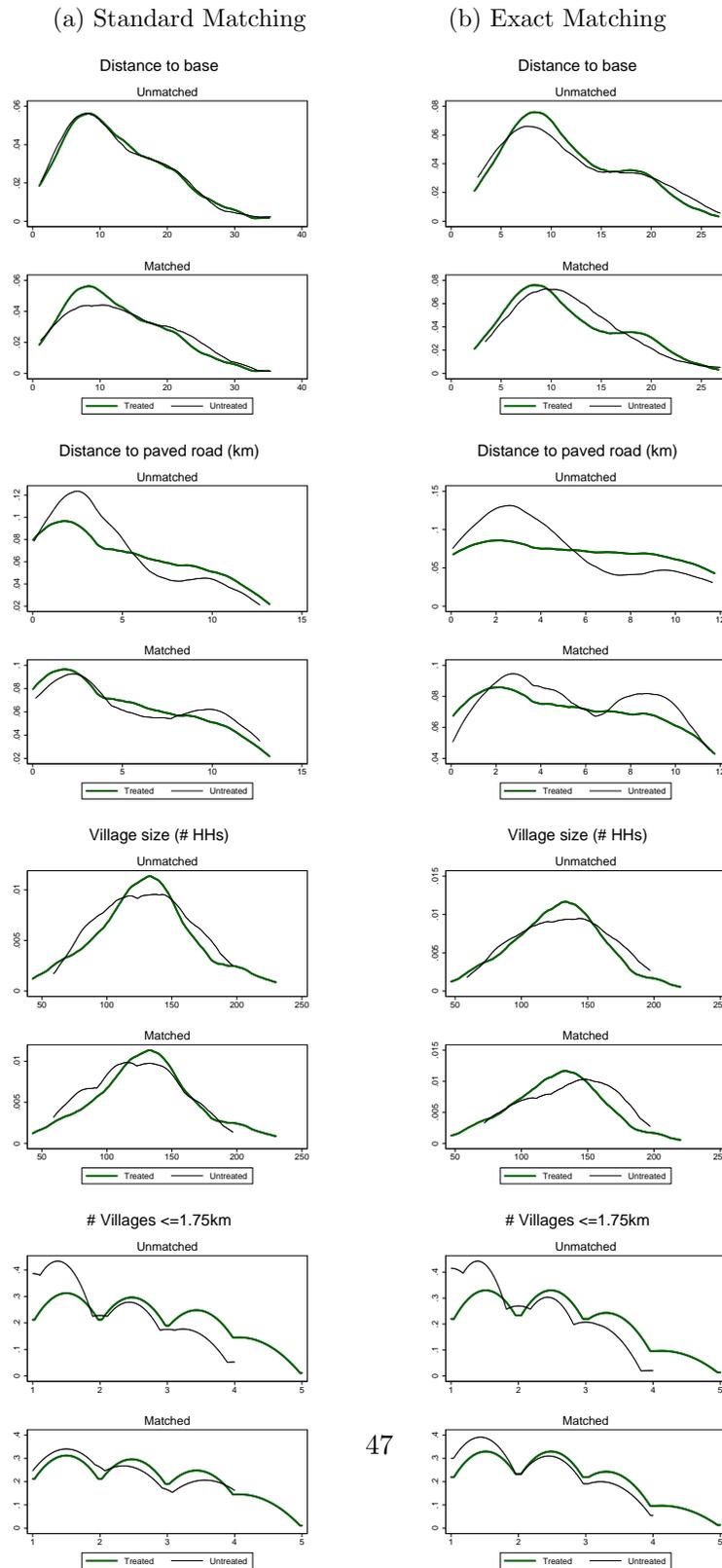
	(1) Exact Matching on Number Villages within 1.75 km
Contribution	33.935*** (11.568)
Observations	118

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 densities with only 0 or 1 village in either treatment or control groups.

#### 5.5.4 Balance

Figure 10 presents graphically the distributions of matching covariates by treatment status, in order to evaluate balance of the matching strategy. For the most part 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 these differences are not very substantial.

Figure 10: Balance of Matching Covariates



Density of covariates in matching estimations by treatment status, before and after matching. (a)  $N = 117$  (b)  $N = 101$  observations.

### 5.5.5 Alternative definitions of CCs

Table 14 presents the effect of CCs for different classifications. Column (1) replicates the main results in the paper, defining an individual as a CC if they contribute in round 2 the mode of contributions in round 1. Column (2) uses the median, while column (3) uses the mean. From the table, one can see that mode and median results are consistent, while the mean results are not significant. However, the coefficient on the interaction for the mean definition is of the same magnitude, though the estimates are noisier. This could suggest that participants found it easier to gauge the mode or median given the experimental protocol, which involved contributions being counted one by one. Additionally, the definition using mode or median appear to better fit the data than the mean.

Table 14: Robustness: Different definitions of CCs

	(1) Mode	(2) Median	(3) Mean
Treatment Status	-10.246 (15.550)	-18.048 (17.118)	-4.613 (19.818)
Conditional Cooperator (CC)	88.166*** (25.901)	95.128*** (31.976)	-16.900 (61.003)
CC $\times$ Treatment	88.669*** (31.976)	92.109** (42.944)	110.617 (71.243)
Distance to base (km)	-3.349 (2.558)	-3.614 (2.760)	-2.359 (3.098)
Distance to paved road (km)	-0.680 (3.862)	-0.956 (3.918)	-1.821 (4.440)
Village Size (# HHs)	0.088 (0.114)	-0.069 (0.120)	0.055 (0.132)
# Villages $\leq$ 1.75 km	0.672 (4.052)	-0.308 (4.512)	0.928 (5.540)
Years of Education	-6.332 (5.075)	-1.875 (5.416)	-3.871 (6.728)
Female	33.779 (33.382)	26.369 (39.261)	81.995* (45.921)
Age	4.352 (18.510)	13.983 (18.949)	10.391 (23.840)
Controls	✓	✓	✓
Sector Fixed Effects	✓	✓	✓
$R^2$	0.58	0.53	0.35
Observations	146	146	146

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at \* 0.1; \*\* 0.05; \*\*\* 0.01. Robust standard errors in parentheses. Specifications follow those in Table 5 column 3.

### 5.5.6 Conditional Cooperation and Heterogeneous Effects

Table 15 presents further analysis of potential heterogeneous effects, using OLS. Online Appendix H presents the analogous table using matching strategies, though power is limited. Villages are split according to the median value of the variable of interest, and average treatment effects are estimated for this sub-sample analogous to Column (3) of Table 5. Chow Test statistics examine whether the differences across these sub-samples are significant. Along the different dimensions, only for CCs is the difference in the two sub-samples statistically significantly different (at the 10% level).

Table 15: Heterogeneous Effects: Past participating neighbors within 1.75 km

	(1) OLS	(2) N
By Conditionally Cooperative		
More Conditionally Cooperative	39.027* (20.929)	58
Less Conditionally Cooperative	-2.476 (16.369)	71
P-Value (Chow test)	[ 0.056]*	
By Age		
Older	39.729* (23.872)	73
Younger	45.552** (20.785)	71
P-Value (Chow test)	[ 0.780]	
By Education		
More Educated	43.306** (20.777)	72
Less Educated	18.653 (23.098)	69
P-Value (Chow test)	[ 0.357]	
By Number Others Known		
More People Known	37.394* (21.665)	70
Less People Known	38.522* (22.809)	72
P-Value (Chow test)	[ 0.770]	

OLS regression of specification (3) in Table 5. Dependent variable is contributions. Selected covariates split by median value. Significantly different from zero at \* 0.1; \*\* 0.05; \*\*\* 0.01. Robust Standard Errors in parentheses.

## References

- Avdeenko, Alexandra and Michael J Gilligan**, “International Interventions to Build Social Capital: Evidence from a Field Experiment in Sudan,” *American Political Science Review*, 2015, pp. 1–45.
- Baldassarri, Delia**, “Cooperative Networks: Altruism, Group Solidarity, Reciprocity, and Sanctioning in Ugandan Producer Organizations,” *American Journal of Sociology*, sep 2015, *121* (2), 355–395.
- Banerjee, Abhijit V., Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, and Michael Walton**, “From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application,” *Journal of Economic Perspectives*, 2017, *31* (4), 73–102.
- Bednar, Jenna, Yan Chen, Tracy Xiao Liu, and Scott Page**, “Behavioral spillovers and cognitive load in multiple games: An experimental study,” *Games and Economic Behavior*, jan 2012, *74* (1), 12–31.
- 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.
- Bernal, Adriana, Juan-Camilo Cárdenas, Laia Domenech, Ruth Meinzen-Dick, and Sarmiento Paula J.**, “Social Learning through Economic Games in the Field,” *mimeo*, 2016, pp. 1–33.
- Bochet, Olivier and Louis Putterman**, “Not just babble: Opening the black box of communication in a voluntary contribution experiment,” *European Economic Review*, 2009, *53* (3), 309–326.
- Brosig, Jeannette, J Weimann, Axel Ockenfels, and Joachim Weinmann**, “The effect of communication media on cooperation,” *German Economic Review*, 2003, *4* (2), 217–241.
- Cardenas, Juan Camilo and Jeffrey P. Carpenter**, “Three themes on field experiments and economic development,” in John A. List Glenn W. Harrison, Jeffrey Carpenter, ed., *Field Experiments in Economics*, 2005, pp. 71–123.

- Casaburi, Lorenzo and Rocco Macchiavello**, “Demand and Supply of Infrequent Payments as a Commitment Device: Evidence from Kenya,” *American Economic Review*, 2018.
- Chaudhuri, Ananish**, “Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature,” *Experimental Economics*, 2011, 14, 47–83.
- , **Sara Graziano, and Pushkar Maitra**, “Social Learning and Norms in a Public Goods Experiment with Inter-Generational Advice,” *Review of Economic Studies*, apr 2006, 73 (2), 357–380.
- Crawford, Vincent**, “A Survey of Experiments on Communication via Cheap Talk,” *Journal of Economic Theory*, 1998, 78 (2), 286–298.
- Duflo, Esther**, “Field experiments in development economics,” in “Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress, Volume II” 2009.
- **and Emmanuel Saez**, “Participation and investment decisions in a retirement plan: The influence of colleagues’ choices,” *Journal of Public Economics*, 2002, 85 (1), 121–148.
- Dufwenberg, Martin, Simon Gächter, and Heike Hennig-Schmidt**, “The framing of games and the psychology of play,” *Games and Economic Behavior*, 2011, 73 (2), 459–478.
- Dupas, Pascaline**, “Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment,” *Econometrica*, 2014, 82 (1), 197–228.
- Enos, Ryan D. and Noam Gidron**, “Exclusion and Cooperation in Diverse Societies: Experimental Evidence from Israel,” *American Political Science Review*, jul 2018, pp. 1–16.
- Giné, Xavier, Pamela Jakiela, Dean Karlan, and Jonathan Morduch**, “Microfinance games,” *American Economic Journal: Applied Economics*, 2010.
- Gneezy, Uri and Alex Imas**, “Lab in the field: Measuring preferences in the wild,” in “Handbook of Field Experiments,” Cambridge: Cambridge University Press, 2017, pp. 1–40.

- Harrison, Glenn W and John a List**, “Field Experiments,” *Journal of Economics Literature*, 2004, *42* (4), 1009–1055.
- Hartig, Björn, Bernd Irlenbusch, and Felix Kölle**, “Conditioning on what? Heterogeneous contributions and conditional cooperation,” *Journal of Behavioral and Experimental Economics*, apr 2015, *55*, 48–64.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath**, “In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies,” *American Economic Review*, may 2001, *91* (2), 73–78.
- Imbens, Guido W. and Donald B. Rubin**, *Causal inference: For statistics, social, and biomedical sciences an introduction* 2015.
- Isaac, R. Mark and James M. Walker**, “Communication and free-riding behavior: the voluntary contributions mechanism,” *Economic Inquiry*, 1988, *26* (4), 585–608.
- **and James M Walker**, “Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism,” *The Quarterly Journal of Economics*, 1988, *103* (1), 179–199.
- Jakiela, Pamela and Owen Ozier**, “Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies,” *Review of Economic Studies*, 2015.
- Kosfeld, Michael and Devesh Rustagi**, “Leader Punishment and Cooperation in Groups: Experimental Field Evidence from Commons Management in Ethiopia †,” *American Economic Review*, 2015, *105* (2), 747–783.
- Ledyard, John O.**, “Public Goods: A Survey of Experimental Research,” *Social Science*, jan 1995, *35* (12), 111–194.
- List, John A.**, “Field Experiments: A Bridge between Lab and Naturally Occurring Data,” *The B.E. Journal of Economic Analysis & Policy*, 2007.
- Miguel, Edward and Michael Kremer**, “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities,” *Econometrica*, 2004, *72* (1), 159–217.

- Rosenbaum, Paul R and Donald B Rubin**, “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, apr 1983, *70* (1), 41.
- Sally, David**, “Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992,” *Rationality and Society*, jan 1995, *7* (1), 58–92.
- Schotter, Andrew**, “Decision making with naive advice,” *American Economic Review*, 2003, *93* (2), 196–201.
- Sinharoy, Sheela S., Wolf-Peter Schmidt, Kris Cox, Zachary Clemence, Leodomir Mfura, Ronald Wendt, Sophie Boisson, Erin Crossett, Karen A. Grépin, William Jack, Jeanine Condo, James Habyarimana, and Thomas Clasen**, “Child diarrhoea and nutritional status in rural Rwanda: a cross-sectional study to explore contributing environmental and demographic factors,” *Tropical Medicine & International Health*, aug 2016, *21* (8), 956–964.
- Sinharoy, Sheela S, Wolf-Peter Schmidt, Ronald Wendt, Leodomir Mfura, Erin Crossett, Karen A. Grépin, William Jack, Bernard Ngabo Rwabufigiri, James Habyarimana, and Thomas Clasen**, “Effect of community health clubs on child diarrhoea in western Rwanda: cluster-randomised controlled trial,” *The Lancet Global Health*, jul 2017, *5* (7), e699–e709.
- Viceisza, A. C. G.**, “Treating the field as a lab: A basic guide to conducting economics experiments for policymaking,” in “Food Security in Practice Technical Guide,” Washington, D.C.: International Food Policy Research Institute, 2012.
- Zizzo, Daniel John**, “Experimenter demand effects in economic experiments,” *Experimental Economics*, mar 2010, *13* (1), 75–98.

Supplementary Online Material for:  
Identifying communication spillovers in lab in the field  
experiments

Alexander Coutts\*  
Nova School of Business and Economics

October, 2019

**A Results for Distance of 2 km**

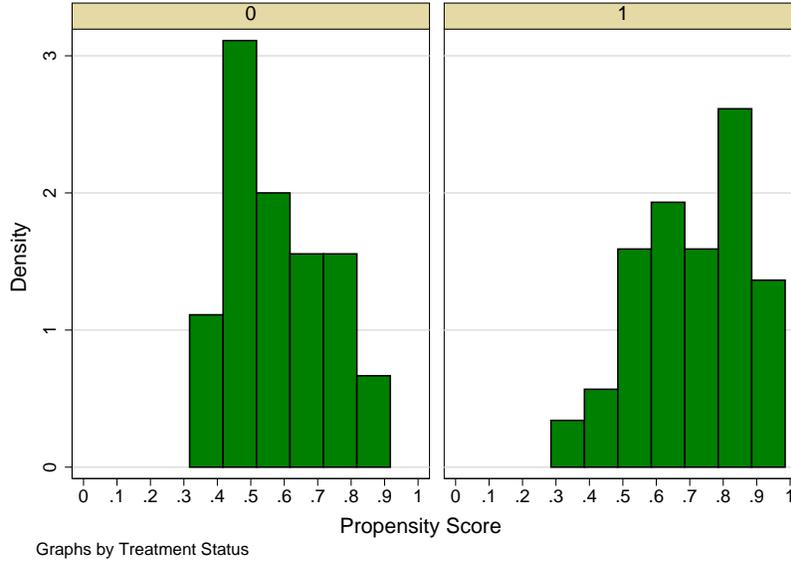
This section replicates the central results of the analysis in the main paper for the distance of 2 km rather than 1.75 km. Results are highly consistent, across both distances.

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

	Treatment	Control	Difference
<i>Available to planner</i>			
Distance to base (km)	11.91	15.02	-3.10**
Village Size (# HHs)	131.32	132.24	-0.92
# Villages $\leq$ 2 km	3.38	1.84	1.54***
Distance to paved road (km)	4.60	4.93	-0.33
<i>Unavailable to planner</i>			
Average Contribution	266.50	237.15	29.35***
Proportion Female	0.74	0.74	0.00
Average Age	35.27	35.56	-0.29
Average Years of Education	4.49	4.44	0.05
Community Cooperation Index	0.81	0.82	-0.00
Community Effort Index	0.79	0.79	-0.00
Average Number current Participants Known	2.51	2.64	-0.14
General Trust Index	0.73	0.73	-0.00
Observations	89	58	147

\*Nova School of Business and Economics, Faculdade de Economia, Campus de Carcavelos, Rua da Holanda, n.1, 2775-405 Carcavelos, Portugal; alexander.coutts@novasbe.pt

Figure A1: Propensity Scores by Treatment Status



Distribution of propensity score.  $N = 134$ .

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

	(1) Standard Matching	(2) Exact Matching
Contribution	36.642*** (11.569)	36.700** (16.038)
Observations	112	88

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	29.355*** (10.518)	33.085*** (10.992)	31.483** (13.075)
Distance to base (km)		-0.067 (0.862)	-0.392 (3.224)
Distance to paved road (km)		4.108* (1.824)	0.570 (4.617)
Village Size (# HHs)		-0.031*** (0.113)	0.065*** (0.141)
# Villages $\leq$ 2 km		-2.396 (4.474)	0.443 (5.354)
Years of Education		1.824 (5.401)	-0.997 (6.815)
Female		94.355** (39.533)	79.481* (43.673)
Age		9.105 (26.208)	13.268 (25.072)
Age <sup>2</sup>		-0.107*** (0.363)	-0.170*** (0.351)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.05	0.20	0.29
Observations	147	146	146

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 3 in main paper.

Table A4: The Role of Conditional Cooperators (CCs)

	(1)	(2)	(3)
Treatment Status	-0.858 (13.864)	-1.685 (15.060)	-10.123 (17.091)
Conditional Cooperator (CC)	108.026*** (34.121)	87.630** (35.596)	80.114** (32.640)
CC × Treatment	59.178 (37.493)	71.191* (38.725)	84.818** (37.455)
Distance to base (km)		-0.242* (0.685)	-2.425 (2.593)
Distance to paved road (km)		1.415 (1.439)	0.328 (3.938)
Village Size (# HHs)		-0.028*** (0.102)	0.071*** (0.115)
(mean) popdensity2000		-2.657 (3.223)	0.718 (3.557)
Years of Education		-1.956 (4.643)	-4.184 (5.108)
Female		67.818** (32.076)	42.670 (32.869)
Age		-0.328 (17.671)	7.995 (18.003)
Age <sup>2</sup>		0.014*** (0.245)	-0.100*** (0.253)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.41	0.48	0.56
Observations	147	146	146

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at \* 0.1; \*\* 0.05; \*\*\* 0.01. Robust standard errors in parentheses. Treatment defined as having past participating neighbors within 2 km. Controls includes all remaining variables found in Table 3 in main paper.

## B 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 B1 one can see that from this small set of variables, there are no significant differences across the villages assigned to treatment versus control.

Table B1: Balance of treatment: Past participating neighbors within 1.75 km (Additional Variables)

	Treatment	Control	Difference
<i>Unavailable to planner</i>			
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.

## C 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 C1 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 C1 shows that the order of visit within sector is associated with significantly higher contributions.

The next Table C2 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 D1 shows that the total number of villages in the sector is not driving this result.

Finally Table C3 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<sup>th</sup> percentile, 5.17km.<sup>1</sup> From column 3, one can see that for every additional kilometer to the nearest participating village, contributions are reduced by 9.6, or about 4% of the average.

---

<sup>1</sup>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 C1: Effect of Ranking of Order Visited Within Sector

	(1)	(2)	(3)
Rank of Visit (within Sector)	16.214*** (5.125)	15.192*** (4.828)	16.469*** (6.218)
Distance to base (km)		0.169 (0.893)	-1.875 (3.006)
Distance to paved road (km)		3.097* (1.852)	-1.471 (4.455)
Village Size (# HHs)		-0.004 (0.116)	0.065 (0.138)
# Villages $\leq$ 1.75 km		4.761 (4.540)	8.840* (4.979)
Years of Education		0.686 (5.320)	-0.870 (7.229)
Female		97.334** (39.712)	79.427* (42.347)
Age		14.246 (25.280)	18.887 (25.595)
Age <sup>2</sup>		-0.177 (0.348)	-0.248 (0.357)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.06	0.20	0.30
Observations	147	146	146

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 3 in main paper.

Table C2: Effect of Number of Previous Participating Villages in Sector

	(1)	(2)	(3)
# Villages before in Sector	12.351** (4.903)	12.796** (5.038)	13.241** (5.976)
Distance to base (km)		0.322 (0.888)	-2.083 (3.065)
Distance to paved road (km)		3.960** (1.866)	-1.113 (4.459)
Village Size (# HHs)		-0.022 (0.116)	0.053 (0.139)
# Villages $\leq$ 1.75 km		5.374 (4.539)	9.210* (4.944)
Years of Education		1.801 (5.502)	-1.449 (7.126)
Female		94.324** (40.455)	75.439* (42.572)
Age		12.221 (26.502)	16.430 (25.468)
Age <sup>2</sup>		-0.152 (0.364)	-0.216 (0.356)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.04	0.19	0.29
Observations	147	146	146

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 3 in main paper.

Table C3: Effect of Distance to Nearest Past Participant

	(1)	(2)	(3)
Distance to Nearest Past Participant	-9.030** (3.760)	-7.752** (3.864)	-9.606** (4.416)
Distance to base (km)		-0.030 (0.853)	-0.961 (3.234)
Distance to paved road (km)		4.139** (1.880)	1.047 (5.007)
Village Size (# HHs)		-0.050 (0.119)	0.037 (0.146)
# Villages $\leq$ 1.75 km		2.218 (4.491)	5.715 (5.080)
Years of Education		2.888 (5.752)	-1.103 (7.503)
Female		76.432* (41.111)	60.868 (44.294)
Age		8.985 (26.390)	11.765 (25.172)
Age <sup>2</sup>		-0.115 (0.365)	-0.159 (0.351)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.03	0.16	0.26
Observations	144	143	143

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 A1. Distance to nearest past participant is in kilometers, and is top-coded at the 95<sup>th</sup> percentile (5.17km).

## D Placebo Tests for Measuring Opportunities for Communication

This section presents three placebo tests for measuring opportunities for communication. The previous Table C2 examined the number of villages that were previously visited within the same sector, and found a strong significant relation with contributions. Table D1 examines the total number of villages in the sector, irrespective of the date of visit. One can see that 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 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 D2. While the coefficient is positive it is not significant at conventional levels. 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 D3 presents the same analysis of the village density, but restricting 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, for the control, there is no relationship between village density and contributions. Additionally the coefficient is much smaller in magnitude than Table D2, and sometimes negative.

Table D1: Effect of Total Number of Villages Within Sector

	(1)	(2)
# Villages in Sector	6.194 (4.611)	4.755 (4.982)
Distance to base (km)		0.143 (0.936)
Distance to paved road (km)		4.107** (1.891)
Village Size (# HHs)		-0.015 (0.121)
# Villages $\leq$ 1.75 km		4.696 (4.574)
Years of Education		0.899 (5.589)
Female		84.170** (42.288)
Age		8.923 (26.292)
Age <sup>2</sup>		-0.118 (0.362)
Controls		✓
Sector Fixed Effects		
$R^2$	0.01	0.16
Observations	147	146

Analysis uses OLS regression. Dependent variable is contributions. Significantly different from 0 at \* 0.1; \*\* 0.05; \*\*\* 0.01. Robust standard errors in parentheses. As there is no variation across sectors in number of villages within the sector, the specification with sector fixed effects is omitted. Variable “# Villages in Sector” has been standardized. Controls includes all remaining variables found in Table 3 in main paper.

Table D2: Effect of Number of Neighbors Within 1.75 km

	(1)	(2)	(3)	(4)
# Villages $\leq$ 1.75 km	8.245 (5.903)	6.707 (6.353)	11.358 (6.943)	3.188 (7.873)
Treatment				29.773** (12.032)
Distance to base (km)		-0.146 (0.888)	-1.133 (3.173)	-1.312 (3.188)
Distance to paved road (km)		4.163** (1.897)	-0.354 (4.504)	-0.307 (4.520)
Village Size (# HHs)		-0.024 (0.123)	0.064 (0.143)	0.094 (0.142)
Years of Education		0.308 (5.552)	-3.409 (6.773)	-3.146 (6.881)
Female		84.679** (42.007)	60.287 (42.719)	76.834* (45.056)
Age		8.796 (26.366)	11.910 (25.267)	15.083 (25.444)
Age <sup>2</sup>		-0.115 (0.364)	-0.167 (0.353)	-0.206 (0.354)
Controls		✓	✓	✓
Sector Fixed Effects			✓	✓
$R^2$	0.01	0.15	0.26	0.30
Observations	147	146	146	146

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 3 in main paper.

Table D3: Effect of Number of Neighbors Within 1.75 km (Control Only)

	(1)	(2)	(3)
# Villages $\leq$ 1.75 km	0.960 (6.229)	-3.336 (7.240)	1.749 (8.287)
Distance to base (km)		-1.014 (1.028)	-6.295 (3.888)
Distance to paved road (km)		2.752 (2.406)	-0.379 (4.945)
Village Size (# HHs)		-0.010 (0.118)	0.163 (0.161)
Years of Education		-1.343 (6.350)	-7.064 (6.679)
Female		99.465 (64.834)	79.550 (72.184)
Age		-15.966 (34.924)	-2.218 (45.311)
Age <sup>2</sup>		0.213 (0.486)	0.023 (0.629)
Believes trustworthy		-8.757 (79.463)	-9.098 (78.786)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.00	0.19	0.48
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 3 in main paper.

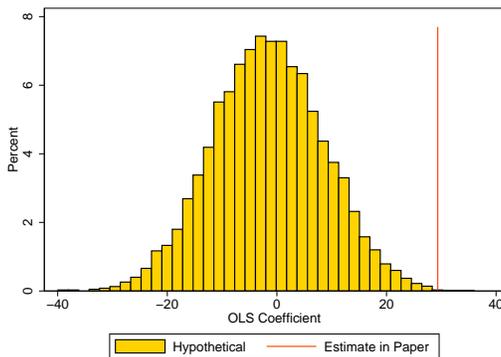
## E Counterfactual Planner and Instrumental Variable Analysis

This section expands on the robustness analysis conducted in Section 3.10 in the main paper. In the first part of that section, I simulated 10,000 random paths, calculated the resulting hypothetical treatment, and re-estimated the main analysis of this paper for each such path. Figure E1 presents the results for the distribution of coefficients on these counterfactual treatments, using the empirical OLS strategy which appears in Column (3) of Table 5 in the main paper. As can be seen, the estimates in the main paper exceed 99.94% of the estimated coefficients in these counterfactual treatments.

Figures E2 presents the analogous figures for the (a) standard and (b) exact matching strategies, respectively, analogous to the results presented in Table 4 of the main paper. From these figures, one can see that the estimates in the main paper exceed 99.87% of the estimated ATEs for the hypothetical treatments for standard matching, and 100% of the estimated ATEs for exact matching on sector.

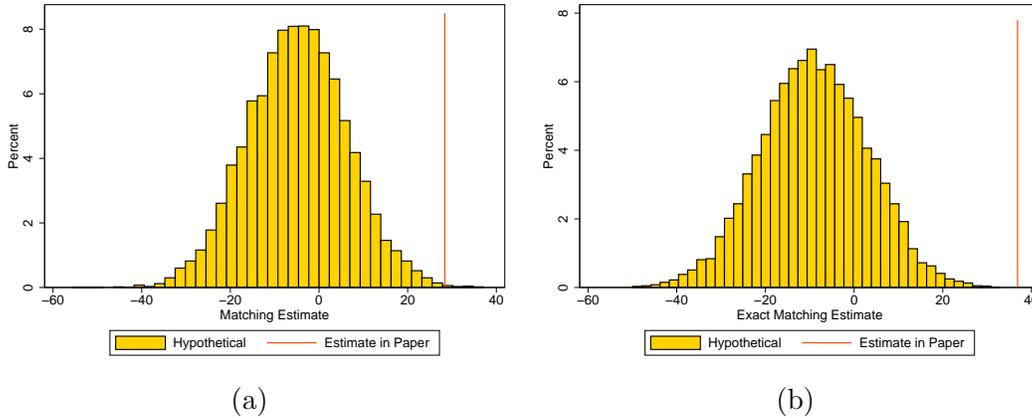
One note is that the average counterfactual estimate of all three specifications is negative. This implies that within the sample of villages, those that are more likely to be treated, i.e. those with more neighbors within 1.75 km, actually contribute less on average. Overall these simulations suggest that the results of the main paper would be exceptionally rare to occur by chance. Given the unlikeliness of observing large positive estimates, these results also reinforce that the estimates observed are likely to be due to communication spillovers. The alternative explanation would be that the planner selected a specific order which generated exceptionally large differences in average contributions between the treatment and control villages as defined in this paper.

Figure E1: Distribution of OLS Coefficients in Counterfactual Treatments



Distribution of 10,000 estimated coefficients from the main OLS specification of the paper, using counterfactual treatments generated from 10,000 random orderings of villages.

Figure E2: Distribution of Matching Estimates in Counterfactual Treatments



Distribution of 10,000 estimated ATE estimates from the (a) standard matching and (b) exact matching specifications of the paper, using counterfactual treatments generated from 10,000 random orderings of villages.

Next I examine the instrumental variables robustness analysis. To generate an instrument that is exogenous to any potential private information of the planner, I simulate six hypothetical routes, using three observables available to the planner, in increasing and decreasing order. Village density within 1.75 km is not utilized, since there is insufficient variation in this variable to generate a deterministic ordering. 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.<sup>2</sup>

After generating the order, the resulting instrument is calculated as the number of hypothetical past participating neighbors (within a 1.75 km radius). Thus, 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). According to this method the instruments selected are those which have ordered sectors and villages by (1) the distance to base (nearest to farthest); (2) the distance to base (farthest to nearest); and (3) the

<sup>2</sup>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.

distance to a paved road (nearest to farthest).

Table E1 presents the 2SLS regressions using identical co-variates as the main OLS analysis, Table 5, in the paper. It is clear from the low F-statistics and large standard errors on treatment status in the latter two columns that the instruments appear weak. The reason for the large difference between the first and other columns is a high correlation between the instruments used and the density of villages within 1.75 km, ranging from  $\rho = 0.59$  to  $0.68$ . This is not surprising, as discussed in the paper, there is a mechanical relationship between density and the instrumental variable treatments derived from hypothetical orderings. In Table 8 in the main paper, to avoid this collinearity between the density and the instruments, a larger density is utilised, 3 km. These two density variables are highly correlated ( $\rho = 0.58$ ), but the 3 km density is less so with the instruments (ranging from  $\rho = 0.28$  to  $0.45$ ).

Table E1: IV-2SLS - Effect of Presence of Past Participating Neighbors Within 1.75 km

	(1)	(2)	(3)
Treatment Status	44.376*	48.462	86.622
	(24.175)	(83.556)	(56.175)
Distance to base (km)		-0.125	-1.654
		(0.801)	(3.216)
Distance to paved road (km)		3.547*	-0.218
		(1.843)	(4.391)
Village Size (# HHs)		0.014	0.151
		(0.127)	(0.145)
# Villages $\leq 1.75$ km		-4.082	-8.950
		(16.083)	(12.500)
Years of Education		0.263	-2.645
		(5.282)	(6.794)
Female		95.939**	108.428*
		(45.249)	(56.091)
Age		11.552	21.141
		(24.976)	(24.703)
Age <sup>2</sup>		-0.146	-0.280
		(0.347)	(0.341)
Controls		✓	✓
Sector Fixed Effects			✓
First stage F-Statistic	25.31	0.66	3.99
$R^2$	0.04	0.18	0.17
Observations	147	146	146

Analysis uses instrumental variable two-stage least squares (IV-2SLS) regression. Dependent variable is contributions. Treatment Status is instrumented with three instruments for the number of past participants calculated from hypothetical routes, see description in text for details. Significantly different from 0 at \* 0.1; \*\* 0.05; \*\*\* 0.01. Robust standard errors in parentheses. Controls includes all remaining variables found in Table A1.

## F Examining Potential Changes in Cooperation Norms

Table F1 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 F1: Effect of Presence of Past Participating Neighbors Within 1.75 km on Cooperation Norms

	(1)	(2)	(3)
Treatment Status	-0.014 (0.022)	-0.013 (0.021)	-0.006 (0.020)
Distance to base (km)		-0.001 (0.001)	0.001 (0.006)
Distance to paved road (km)		0.001 (0.002)	-0.008 (0.007)
Village Size (# HHs)		0.000 (0.000)	0.000 (0.000)
# Villages $\leq$ 1.75 km		0.012 (0.009)	0.013 (0.009)
Years of Education		-0.004 (0.008)	0.006 (0.009)
Female		-0.019 (0.066)	0.079 (0.069)
Age		0.055 (0.034)	0.053 (0.034)
Age <sup>2</sup>		-0.001* (0.000)	-0.001* (0.000)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.00	0.46	0.59
Observations	147	146	146

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 3 in main paper.

## G Restricting OLS Estimation of Conditional Cooperation

In this section Table G1 presents the analog to Table 8 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 G1 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 G1: The Role of Conditional Cooperators (CCs)

	(1)	(2)	(3)
Treatment Status	-18.281 (20.923)	-23.994 (20.319)	-39.145* (23.602)
Conditional Cooperator (CC)	95.670* (50.282)	62.762 (54.585)	112.948** (48.505)
CC $\times$ Treatment	107.307** (53.307)	127.676** (59.274)	116.569** (54.463)
Distance to base (km)		-0.912 (0.826)	-8.751** (3.534)
Distance to paved road (km)		1.166 (1.694)	4.015 (5.944)
Village Size (# HHs)		-0.071 (0.117)	-0.096 (0.127)
# Villages $\leq$ 1.75 km		-1.425 (6.219)	1.220 (7.538)
Years of Education		-8.875 (6.532)	-17.755** (8.726)
Female		26.368 (48.430)	-8.184 (50.122)
Age		1.876 (28.416)	3.032 (27.102)
Age <sup>2</sup>		-0.071 (0.396)	-0.082 (0.372)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.51	0.62	0.77
Observations	76	76	76

Analysis uses OLS regression. Dependent variable is contributions. Data restricted to only villages which participated in the same version of the public goods game in the second round. Significantly different from 0 at \* 0.1; \*\* 0.05; \*\*\* 0.01. Robust standard errors in parentheses. Treatment defined as having past participating neighbors within 1.75 km. Controls includes all remaining variables found in Table 3.

## H Heterogeneous Effects with Matching Strategy

Table H1 presents the analogous examination to 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 52 and 57) 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 19 to 25. Thus the results in Column (2) are presented only for consistency, it is not possible to draw any conclusions from this data.

Table H1: Average Effect: Heterogeneous Effects

	(1) Standard Matching	(2) Exact Matching
By Conditionally Cooperative		
More Conditionally Cooperative	54.635*** (15.633)	6.746 (17.069)
Less Conditionally Cooperative	7.809 (11.821)	-4.207 (9.519)
<i>N</i>	57	19
By Age		
Older	43.453*** (14.260)	28.304* (15.657)
Younger	21.568 (16.515)	22.998 (27.084)
<i>N</i>	56	19
By Education		
More Educated	26.965* (15.915)	84.521*** (15.091)
Less Educated	34.871** (14.801)	20.864 (27.248)
<i>N</i>	57	25
By Number Others Known		
More People Known	25.541 (16.620)	38.030 (36.021)
Less People Known	22.813 (18.009)	21.321 (26.259)
<i>N</i>	53	22

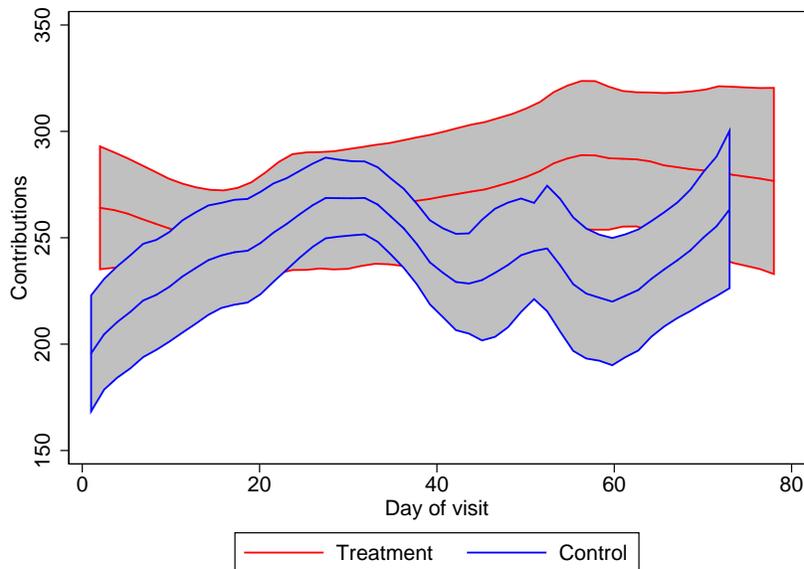
Analysis uses nearest neighbor propensity score matching, with 2 neighbors, with replacement. ATE of treatment on contributions. Significantly different from zero at \* 0.1; \*\* 0.05; \*\*\* 0.01. Abadie-Imbens Robust Standard Errors in parentheses. Values of propensity score outside common support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

## I 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 II presents Epanechnikov kernel-weighted local polynomial smoothing plots regarding the relationship between day 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 alludes to the possibility 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 II: Contributions over time



Epanechnikov kernel-weighted local polynomial smoothing plot showing relationship between day of visit and contributions.

Although visit date and treatment status are not significantly associated, they are positively correlated (F-test p-value 0.121), 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 I1 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 still large and significant.

One can also add date of visit as a further explanatory variable in the OLS regression. Table I2 presents the analogous regression to Table 5 in the main paper, but with the addition of the date of visit. Date of visit is never significant at conventional levels, and does not substantively alter the coefficient on the treatment dummy.

Table I1: ATE of Presence of Past Participating Neighbors Within 1.75 km

	(1) Standard Matching	(2) Exact Matching
Contribution	36.674*** (12.204)	33.511*** (11.227)
Observations	116	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 4, but adds the additional input variable of day of visit.

Table I2: Effect of Presence of Past Participating Neighbors Within 1.75 km

	(1)	(2)	(3)
Treatment Status	28.321*** (10.431)	27.567** (10.757)	26.465** (13.242)
Date of Visit	0.362 (0.237)	0.489 (0.350)	0.701 (1.204)
Distance to base (km)		-0.442 (0.865)	-1.597 (3.236)
Distance to paved road (km)		2.362 (2.072)	-0.370 (4.545)
Village Size (# HHs)		-0.015 (0.119)	0.086 (0.143)
# Villages $\leq$ 1.75 km		-0.135 (4.844)	2.993 (5.758)
Years of Education		-0.177 (5.561)	-2.832 (6.957)
Female		90.165** (39.310)	76.911* (44.563)
Age		13.203 (25.926)	15.880 (25.421)
Age <sup>2</sup>		-0.169 (0.359)	-0.214 (0.355)
Controls		✓	✓
Sector Fixed Effects			✓
$R^2$	0.07	0.21	0.30
Observations	147	146	146

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 3.

## 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.