Collective action in games as in life: Experimental evidence from canal cleaning in Haiti

Abbie Turiansky

September 29, 2015

Job Market Paper[1]

Abstract: When the provision of public goods depends on voluntary contributions, informal institutions and norms can play an important role in increasing contributions. While the literature provides examples of decentralized management of common-pool resources such as irrigation infrastructure, we know little about how successful institutions emerge and evolve. In this paper I ask whether exposure to the strategic considerations of a collective action dilemma in an experimental setting can change behavior in real-world scenarios in which those individuals face similar strategic trade-offs. Among 800 rice farmers who are part of an agricultural technology adoption study in rural Haiti, I randomly selected 300 to participate in public goods games framed to mimic the real trade-off they face between private work and participation in the management of shared canals. Over the subsequent planting season, the local irrigation association organized voluntary canal-cleaning work days to manage the shared canal systems that irrigate farmers' fields. Treated farmers were 66% more likely than the control group to volunteer. The mechanism through which the experiments seem to operate is by affecting participants' expectations of others' contributions to the public good, suggesting that experiments provide a setting in which to learn about one's neighbors and develop common norms of behavior.

[1]I thank Travis Lybbert, Michael Carter and Katrina Jessoe for their guidance through all stages of the project. Funding for the SRI in Haiti project is provided by the BASIS Assets and Market Access Collaborative Research Support Program. Further funding for field experiments was provided by the Henry A. Jastro Graduate Research Scholarship Fund. Research collaboration and support was provided by the Faculte d'Agronomie et Medecin Veterinaire in Haiti and Oxfam America. Special thanks go to the data collection team in Haiti, especially Wesner Antoine and Mackenson Desrosiers. I also thank Thomas Barre, Isabel Call, Diane Charlton, Felipe Dizon, Jon Einar Flatnes, Quentin Stoeffler, Emilia Tjernstrom, Asare Twum-Barima, and Eliana Zeballos for helpful comments and discussions. All errors are my own. Contact: Abbie Turiansky, Department of Agricultural and Resource Economics, University of California, Davis; aturiansky(at)ucdavis.edu.

1 Introduction

The important role that local institutions, formal and informal, play in shaping development outcomes is widely recognized. A growing literature links social capital and norms to a range of development challenges, from management of shared resources to group lending to sanitation behavior. A substantial literature has analyzed the characteristics of successful decentralized management of common-pool resources such as irrigation systems (e.g. Ostrom (1990); Cárdenas, Rodriguez, and Johnson (2011); McCarthy, Sadoulet, and de Janvry (2001); Baland and Platteau (1998)). These characteristics include formal local management institutions with rules and enforcement mechanisms as well as informal institutions that include social norms around cooperation and resource use. While the literature has improved our understanding of the types of norms and institutions that are correlated with successful resource management, the study of how to encourage and strengthen such institutions is complicated by the lack of exogenous variation in drivers of institution formation. Improving our understanding of how norms and institutions form and evolve is of great policy relevance as interventions and development policies aim to encourage collective action through strengthening local institutions.

In this paper I ask whether exposure to a collective action problem in an experimental setting can change real-world behavior in a scenario in which the same strategic considerations are relevant. I randomly assign participants to play public goods games to study whether such exposure may shift farmers' beliefs or understanding of a local public goods provision problem and ultimately move real-world behavior toward more socially optimal levels. The use of a lab-in-the-field experiment as a source of exogenous variation provides a unique contribution to the literature on decentralized management of public goods as well as the experimental and behavioral economics literature by demonstrating that experiments can shift beliefs and lead to changes in individual contributions to a public good.

I find that farmers who participated in a public goods game, framed to replicate the coordination problem they face managing shared irrigation infrastructure, were 66% more likely to contribute to canal management during the subsequent planting season. This finding suggests that such experiments may be useful not only to study in-game behavior as an outcome of interest but also for shifting behavior outside of the games. I further find suggestive evidence that one mechanism through which the experiments operate is through an effect on participants' expectations of others' contributions to the public good. Experiments provide an opportunity for participants to learn about the social and technological context in which the public good is provided, and this context includes the behavior one can expect of other beneficiaries of the public good. The increase in real-world public goods contributions primarily among participants who learn that their neighbors are likely to contribute suggests a desire to conform to a norm of contribution levels, and that experiments provide a setting through which individuals can learn about and develop common norms.

I conduct the games with rice farmers in Haiti who depend on a shared irrigation system that lacks a formal management system. While the local irrigation association organizes farmers to contribute to management and cleaning, farmers report that participation is inadequate and that poorly functioning canals lower yields and increase flooding risk. This research context is a classic example of the tragedy of the commons (Hardin (1968)), in which theory predicts that individuals acting in an uncoordinated manner will fail to reach a socially optimal level of the public good. Yet we see examples in a variety of settings where individuals have overcome a collective action dilemma to establish norms, institutions, and rules to achieve a solution that is closer to the social optimum. For example, people everywhere vote, contribute to charities, and organize themselves to establish rules and norms of behavior for mutual benefit. There is an extensive literature examining local institutions that have developed to manage shared resources and provide public goods (see Hess (1999) for an overview).

While the literature provides us with examples and characteristics of local institutions capable of managing resources and providing public goods, we also know that such coordinated behavior fails in many settings. From a policy perspective, understanding how social institutions develop and evolve, and why we see such success in some settings and not in others could provide valuable guidance for encouraging the formation and strengthening of local institutions and norms. Voors and Bulte (2014) suggest two primary pathways for institutional change: (1) change by design, whereby institutions are shaped by intervening outsiders or by strategic decisions of elites, or (2) evolutionary change, in which systems of rules and norms emerge as a result of the uncoordinated choices of many agents. In the latter, individuals' expectations and beliefs change, either as a result of outside intervention or from some internal changes in beliefs, and these new beliefs and expectations change behavioral norms over time, gradually changing rules and constraints.

Many recent interventions aim to create or strengthen local institutions and governance and to involve communities in the development process, but most studies fail to find that outside intervention can successfully change institutional quality through a top-down approach. Casey, Glennerster, and Miguel (2012) find that a program specifically targeting institutional quality had no impact on measures of fundamental social change, such as collective action capacity and inclusion of marginalized groups. Fearon, Humphreys, and Weinstein (2009) similarly find some impact of a Community-Driven Development project on leaders' mobilization efforts for community projects, but no evidence for greater participation by community members. Similar findings of other interventions are reported in Beath, Christia, and Enikolopov (2013), Avdeenko and Gilligan (2014), Mansuri and Rao (2012). In contrast, in their study of post-conflict institutional quality in Burundi, Voors and Bulte (2014) find that changes in institutional quality seem to be explained by changes in beliefs and expectations rather than by outside design. This finding raises the question of how beliefs and norms change, and whether they can be influenced by an external intervention.

Studies have appealed to norms and social comparisons in an attempt to change behavior in a range of contexts, from smoking cessation to school performance to charitable giving. Many of these studies have found that such non-pecuniary strategies can lead to more pro-social behavior even in the absence of substantial personal benefit. For example, Ferraro and Price (2013) find that messages appealing to social norms, and in particular, comparisons with neighbors' behavior, reduce households' water use, at least in the short run, and Allcott and Rogers (2014) find a similar result with respect to energy consumption.

The particular mechanism I use to influence individual behavior is the use of experimental games, framed to replicate the real situation in which cooperation is failing. A large number of laboratory experiments have demonstrated that individuals tend to contribute to public goods even when the rational prediction would be to participate nothing. In these experiments, researchers repeatedly find factors that should not change the theoretically predicted behavior, such as expectations of others' behavior, the ability to communicate, and the number of rounds played, correlate strongly and consistently with actual contributions (see Ledyard (1995) for an overview). There is some evidence that participants in laboratory and field experiments exhibit signs of learning over the course of repeated games, changing behavior over repetitions and even into similar games (Bó and Fréchette (2011); Duffy and Ochs (2009)). Experiments can expose participants to the strategic considerations of certain scenarios and prime individuals to behave differently in real-world scenarios in which the same considerations are relevant. When experiments are framed to be relevant to farmers' experiences, the experiments may make salient the importance of cooperative behavior, which may, in turn, lead to stronger resource management institutions.

The outcome of interest in this paper is participation in the cleaning and maintenance of the shared canal system, a public good that benefits all farmers whose plots are watered by the system. Experimental economics has traditionally focused on testing theories within controlled experiments, in which in-game behavior is the outcome of interest. Several recent studies (Karlan (2005); Hoff and Pandey (2006); Carter and Castillo (2011)) have expanded the field to link behavioral measures elicited from field experiments to economic outcomes, demonstrating the predictive power of behavioral experiments. Karlan (2005) finds that players deemed more trustworthy based on experimental measures repay loans at a higher rate and save more money, while those deemed more trusting save less and are more likely to drop out of a credit association. Several earlier studies conduct field experiments and link behavior from those experiments to community- and individual-level characteristics such as bargaining behavior, market integration, and community cohesion (Roth et al. (1991); Henrich et al. (2001); Barr (2003)). These studies provide evidence that behavioral measures from experiments can reliably predict real-world behavior in some settings.

The use of in-game behavior as an outcome in itself or as a predictor of real-world behavior implies that traits such as trustworthiness and altruism are individual characteristics and that the experiments serve only as a measurement tool. However, if experiments can provide participants with an opportunity to reflect on a problem they face in real life, they may serve as a pedagogical tool as well. Only one study of which I am aware has shown behavioral change as a result of participation in a field experiment. Cardenas and Carpenter (2005) conducted a series of common-pool resource experiments with villagers who rely on a local common-pool resource, and returned to the same villages between six and 20 months later to run the same experiments. They found that decisions in favor of a group-oriented outcome were significantly higher in the second round of experiments among both repeat participants and new participants. They have anecdotal evidence that participants talked with their neighbors after the first round, so even those who did not participate in the first year had an opportunity to learn from the experiments through their neighbors. We do not know whether the behavioral change observed within the later experiments spilled over into behavior change with regards to the real common-pool resource on which they rely.

While we traditionally think of surveys and field experiments as measurement tools, with in-game behavior as the outcome of interest, previous evidence has found that being the subject of research can influence real-world behavior. In a study of several health and microlending projects, Zwane et al. (2011) find that being surveyed increases take-up of medical insurance and use of water treatment technologies, and can bias impact estimates. If we disregard the impact of data collection tools on the behavior of those being studied, we risk biasing our results. On the flip side, we can recognize that tools traditionally used for data collection can play a role in intentionally influencing beliefs and changing behavior, as I do in this study. This paper provides new evidence that exposure to a coordination dilemma in the form of an artefactual field experiment can change behavior in a real world situation in which they face the same strategies and trade-offs. In a setting where institutions are weak, changing cooperative norms and encouraging participation in public goods provision may be instrumental in improving outcomes for farmers currently struggling with low availability of public goods. This research provides a possible mechanism for low-cost behavior change in a context where higher levels of cooperation would raise welfare for the farmers involved.

2 Conceptual Framework

In this section I present a simple model to illustrate the choice farmers make to allocate their private resources to a public good. I then use this model to examine the possible mechanisms through which the experimental intervention may change farmers' decisions over how much he allocates to the public good.

2.1 Model

A farmer divides his resources between cultivating his own plot and contributing to the public good. The public good in this context is the performance of the shared canals. When canals are clear of blockages and well graded, they allow for water to enter and leave the field as necessary and reduce the risk of flooding, thus raising average yields for all farmers who share the canal. Canal performance is a function of the aggregate contributions of all farmers. The farmer's private yields on his plot are a function of the resources he allocates to his farm and the performance of the canal.

Each farmer has a fixed endowment, ω_i, that he can allocate between two inputs into farming: a private composite input, x_i, and his individual contributions, z_i, to a public good, g. The public good is a function of $Z = \sum_i z_i$, the total of all farmers' contributions: $g = g(Z) = g(\sum_i z_i)$. Yield is a function of the private input and the performance of the canals: $Y_i = f(x_i, g(Z))$. Yield is increasing and concave in both x_i and $g(Z)$.

I assume in this model that farmers are uncertain about the functional forms specifying the relationships between resources allocated to the public good, performance of the canals, and farm yields. As a result, farmers make their decisions with respect to *expected* yields:

$$E_i[Y_i] = E_i\left[f(x_i, g(Z))\right]$$

The experimental literature has found that both in laboratory experiment settings and in real-world behavior, contributions to public goods are frequently higher than would be predicted based only on the utility gained from the level of the public good (see, e.g., Fischbacher, Gächter, and Fehr (2001); Isaac, Walker, and Williams (1994); Chaudhuri (2011)). In other words, it is possible that individuals gain utility both through the level of a public good (in our case, through its impact on yields) and directly from their own individual contribution level. Andreoni (1990) coined the term "warm glow" to refer to the utility that one obtains from contributing to a common good, due to altruistic preferences or other personal benefit from giving to others. If we assume that an individual's utility is a function of his expected farm yield and the utility he obtains directly from his own contribution to the public good, we can model his utility function as:

$$U_i = U_i(E_i[Y_i], z_i) = U_i\left(E_i[f(x_i, g(z_i + Z_{-i}))], z_i\right)$$

where U_i is concave and weakly increasing in both Y_i and z_i. The farmer chooses his allocation between his personal farm and the public good in the following choice problem:

$$\max_{x_i, z_i} \; U_i\left(E_i[f(x_i, g(z_i + Z_{-i}))], z_i\right) \tag{1}$$

$$\text{s.t.} \quad x_i + z_i = \omega_i$$

Solving the maximization problem yields the result that equates the marginal utility the farmers gains from allocating a unit of his endowment to each input. The marginal utility from x_i is simply through its effect on expected yields, while the marginal utility from z_i is the sum of the direct utility he obtains from giving to the public good and the utility he gains through the of z_i on expected yields.

$$\frac{\partial U_i}{\partial z_i} + \frac{\partial U_i}{\partial EY_i} \left(\frac{\partial EY_i}{\partial g(Z)} \frac{\partial g(Z)}{\partial z_i} \right) = \frac{\partial U_i}{\partial EY_i} \left(\frac{\partial EY_i}{\partial x_i} \right) \tag{2}$$

If an individual obtains no utility from his own level of public goods contribution and cares only about farm yield, the result simplifies to a standard problem of equating the marginal returns to yield of the private and public good:

$$\frac{\partial EY_i}{\partial g(Z)} \frac{\partial g(Z)}{\partial z_i} = \frac{\partial EY_i}{\partial x_i} \tag{3}$$

2.2 Mechanisms for Influencing Contributions

We can use this model to imagine possible pathways through which an individual's contributions to the public good may be influenced through an intervention. I divide these mechanisms into two categories: technical and social or behavioral. One way in which the experiments may change behavior is by changing farmers' understanding of the technical or physical relationships between labor allocation to canal maintenance, canal performance, and expected rice yields. The experiments could also allow farmers to learn something about the farmers with whom they share the canals, leading to a social learning effect that operates through the utility a farmer gains directly from his contribution to the public good.

1. Technical mechanism: The games may change participants' understanding of or attention to either of the production functions:

   (a) The yield production function: $Y_i = f(x_i, g(Z))$. If the farmer's expected gains from the public good increase, for any given level of public good, an increase in $\frac{\partial Y_i}{\partial g(Z)}$ will require an increase in $\frac{\partial Y_i}{\partial x_i}$, which, given the assumption of diminishing marginal returns to x_i, will decrease the level of x_i relative to z_i.

(b) The public goods production function $g(Z)$. If the expected returns to contributions increase, as was the case for an increase in the expected returns to the public good itself, we would observe an increase in an individuals contributions to the public good.

In both of these cases, the perturbation of the equilibrium provided by the experiment is not to change the functions themselves, but rather to shift participants' knowledge of the functions or their salience. A farmer may not understand well the importance of the public good in the production function for agricultural yields, or if he is aware of its importance he may under-weight it in his behavior if he is not actively conscious of the role of the public good in his production function and as a result does not fully understand the importance of coordination with his neighbors. The experiments may change either the participant's understanding of these relationships, or simply increase the salience in his mind. In either case, we would be able to see a change in behavior as a result.

2. Social mechanism: The games may change the utility attached to one's own contribution level, $U_i(z_i)$. If, for a given level of z_i and Z, the marginal utility an individual farmer obtains from increasing his own contribution increases, we would see an allocation of more of his endowment toward the public good. We can modify the utility function to specify the different pathways through which an individual's contributions affect his utility by defining a function $h_i(z_i, \bar{z}, X_{-i})$, where $\bar{z}$ is the average contribution of the other farmers and X_{-i} is a vector of characteristics of all other farmers sharing the canals. X_{-i} could also include variables characterizing farmer j's relationship to farmer i. With this modification, the farmer's utility function is:

$$U_i = U_i(Y_i, z_i) = U_i\left(f(x_i, g(z_i + Z_{-i})), \alpha_i h_i(z_i, \bar{z}, X_{-i})\right)$$

where α is the weight one places on one's own contributions in his utility function. If an intervention changes the utility one obtains from one's contributions we could characterize this as a change the function h_i or a change in the weight α_i. This shift may occur due to one or more of several factors:

(a) *General altruism:* Through a change in the weight of one's contributions to the public good in one's utility function. We could think of this effect as the general importance of civic-mindedness or of contributing to a common good, and represent it mathematically as a change in the weight parameter α_i.

(b) *Expectations of Others:* Through the farmer's expectations of others' average contributions, $\bar{z}$. An individual's direct utility from contributing may be related to a sense of conforming to a common norm. Individuals may not want to deviate too far from the norm, either in the positive or negative direction. A change in expectations of what others in the group are likely to contribute could, in this case, influence one's contributions in the direction of the norm.

(c) *Social Networks and Relationships:* Through a change in how much one cares about the others who benefit from one's actions. Other-regarding preferences depend on who benefits; for example, one may place more weight on the benefit to a neighbor or family member than one places on the benefit to a stranger. Any change in how one sees the others benefitting from one's actions could change how much utility one gets from one's own contributions to the common good. We can represent such a shift as a change in X_{-i}, a flexible parameter that includes characteristics of other farmers as well as those farmers' relationships to farmer i.

3 Sample and Experimental Design

3.1 Context: Research Area and Population

This study is conducted in conjunction with an agricultural development program designed to evaluate the household-level welfare impacts of the System of Rice Intensification (SRI) in Artibonite, Haiti. SRI is a potentially high-yielding, low external input cultivation method that has been shown to generate substantial and persistent yield increases (Stoop, Uphoff, and Kassam (2002); Sinha and Talati (2007)). However, adoption has been lower than expected given its apparent benefits, and substantial disadoption has been observed in some locations (Moser and Barrett (2003); Takahashi and Barrett (2014)). One possible explanation for low adoption of SRI is its reliance on precise water management. Preliminary evidence in Haiti and elsewhere shows that poorly functioning irrigation systems are a substantial constraint to adoption. In the absence of publicly provided infrastructure, the long-term success of the intervention may depend crucially on the ability of farmers to establish a coordinated system to manage shared irrigation infrastructure. A key component of the research program is to investigate ways in which such an institution may be encouraged among farmers.

The study takes place in the Artibonite Valley, the largest rice-producing region in Haiti. Farmers in the study cultivate land irrigated by a canal system managed by a local irrigation users' association, *Association des Irrigants Liancourt Artibonite (AILA)*, chosen based on its conduciveness to the technology and because of the implementing partner's history working with the local irrigation association. The irrigation system is complex, involving canals and drains at multiple scales, with different entities potentially responsible for the management at each scale. The irrigation system includes a large concrete canal and a series of gates connecting the large canal to production blocks. A production block is defined as a set of plots that receive water from a single irrigation gate connected to the main canal. Blocks range in size from dozens to hundreds of hectares and can include between 200 and over 1,000 farmers.

The local governmental agricultural office is responsible for maintaining the large system of concrete drains and canals, while *AILA* is responsible for maintaining the gates that feed the irrigation system within each block. Within a block, water is distributed to and drained from individual plots via a series of secondary, tertiary, and quaternary canals and drains. Most of these canals and drains serve many individual parcels, while farmers whose plots are not directly adjacent to a canal typically build very small personal canals to connect to the shared system. No formal system currently exists for maintaining the canals and drains within the blocks. Individual canals are the responsibility of the farmer whose plot is served by the canal, but shared canals are maintained inconsistently, if at all. Farmers in the study region have provided anecdotal evidence of a previously common system for managing the shared canals. Under this system, farmers would come together in traditional work days known as *kombits* in which all farmers were expected to participate.

Farmers report that currently the tradition of *kombits* has largely disappeared, and they provide several possible explanations. The population in the study region is relatively transient, and many landowners manage their plots from afar, hiring labor and visiting occasionally to work their farms. The absence of a stable population makes it more difficult to organize communal work days, and absentee landowners are not likely to come to their plots to participate. Farmers and local implementing partners have also anecdotally linked the decline in *kombits* to the recent pattern of aid agencies, aiming to help farmers whose yields are dramatically hurt by poor water management, paying farmers to clean canals. Many farmers say they are unwilling to participate in voluntary communal work days because of the possibility that an external agent may pay them to do the very same work.

Kombits can be organized at multiple levels: *AILA* or another farmers' association can organize *kombits* at the block level to clean the secondary canals that run through the entire block, while small groups of farmers can organize at a more local level for a section of canal that affects only a small number of parcels. In this study, I focus on block-level *kombits* organized by *AILA*. At the block level, canal cleanliness is fairly close to a pure public good: all farmers in the block benefit from it, regardless of their own contribution to its maintenance. At a more local level, the benefits of canal cleaning can be much more private. Farmers describe cleaning the portion of a smaller shared canal that runs by their own plot as primarily benefitting themselves, but with positive spillover effects on others: if the portion of the drain downstream from farmer i's plot is clean, it will help his drainage, but to a lesser extent. No local-level *kombits* were organized during the study period, so any canal-cleaning that happened at the local level took the form of individuals cleaning near their own parcels, primarily benefitting their own plot with positive spillovers on their neighbors. Given the frame of our experimental intervention as contributions to a public good, the bloc-level *kombit* better approximates the experimental frame than any canal cleaning done at a more local level. It is possible that any impact the intervention had on participation canal cleaning could impact both levels, so we can view the impact only on the block-level *kombit* participation as a lower bound for the impact.

3.2 Sample and Research Timeline

The study includes 804 farmers who cultivate land in four of the irrigation blocks located within the system managed by *AILA*. Local collaborating partners selected four blocks that they deemed conducive to the implementation of the agricultural technology to be part of the program. Households were randomly selected in equal numbers from each of the four blocks for inclusion in the study. Prior to the implementation of the agricultural technology project, in February-March 2014, all study households completed a household survey on agricultural and other economic activities, household demographics, and other baseline characteristics.

Following the completion of the survey, we conducted public goods experiments (explained in further detail in Section 3.3) with a random subsample of farmers in April 2014. Approximately 300 farmers were invited to take part in the experiment treatment. These farmers were sampled in geographic clusters based on parcel maps created by a local collaborating partner prior to the baseline survey. To create the sample, I randomly selected 15 farmers from each of the four study blocs, for a total of 60. Using the map of all parcels in the study blocs, I identified the four parcels adjacent to or closest to each randomly selected farmer, creating clusters of five neighboring farmers, a total of 300 individuals. Field guides were provided with lists of these groups of five farmers and told to find four of them for each group. This oversampling was necessary to ensure four participants in each group, as some farmers were unavailable to participate for various reasons. Because of the sampling technique, farmers with more parcels in the study area (the median number of parcels is two) were more likely to be selected, so I control for number of parcels in all analyses. I treat non-compliers in the analysis by using both an Intention to Treat (ITT) measure and a Treatment on the Treat (TOT) measure using the invitation to participate as an exogenous instrument for participation.[2]

During the 2014 rainy season, the local irrigation users' association *AILA* organized periodic canal-cleaning work days (*kombits*) between May and August 2014. *Kombits* were organized separately for each production block, with each block organizing between five and 16 days of canal cleaning. Participation was optional but all farmers were encouraged to participate. *AILA* collected names of all participants each day and provided farmers with a small stipend for each day they participated to cover costs of transportation and food for the day.[3]

3.3 Experimental Design

The treatment consisted of a session of public goods (PG) game played with farmers in April 2014. Each farmer participated in one session, which lasted between two and three hours. Participants

[2]For these measures, we have compliers, non-compliers (never takers), and those not offered treatment. While those offered treatment had the option of not complying, i.e. not participating in the experiments, those not offered treatment were not able to participate. As a result, we have no "always takers" in our population.

[3]The stipend offered was lower than the typical daily wage that farmers in our sample would earn for the day, so farmers faced a real trade-off between participating in the *kombit* and real work.

played in groups of four farmers with neighboring parcels because the games were framed around cleaning canals shared with neighbors. Each session consisted of five groups of four. The setup and structure of the game was constant across all groups and sessions.

Farmers played a straightforward PG game in which they were each given ten chips, representing ten days of work, to allocate between a private payoff, framed as off-farm work, and a shared payoff, framed as cleaning shared canals. The private activity yielded a constant payoff, while the public activity yielded a payoff that depended on the aggregate contributions by all group members. If total contributions exceeded a threshold, the shared canals were clean, and everyone in the group received a higher payoff from their rice fields. If total contributions fell short of the threshold, all participants received a lower payoff from farming.

Payoffs were calibrated so that the social optimum was for the group to contribute the threshold amount, but the private payoff created an incentive for each individual to free-ride on the contributions of others. Because the threshold exceeded each individual's initial endowment, no individual could unilaterally contribute enough to ensure the higher rice yields, so coordination was necessary. These experiments were designed to simulate the real-life tradeoff between working for a private benefit and contributing to a public good, where benefits are shared.

In each round of the game participants were asked whether they expected others in the group to give more, less, or the same, on average, than what they contributed to canal cleaning. These responses are used as measures of expected cooperativeness of one's neighbors. These measures were not incentive-compatible: participants were not given any additional points for correctly anticipating their neighbors' behavior.[4] However, farmers did, on average, predict fairly well how their own behavior would compare with their neighbors', so this measure appears to be relatively valid.

Farmers played three different games during sessions. The first was a single shot game with no communication, which can be thought of as a measure of initial propensity to cooperate in the provision of a public good. In the second game, farmers played repeated, independent rounds of the same game, presented as multiple rice seasons, with the ability to communicate between each round. Communication was verbal: each group was given approximately one minute to discuss

[4] Pre-tests showed that such an award system was difficult for farmers to understand and drew too much focus away from the primary decision-making in the game.

anything they wanted before making their decisions in private, though they were instructed not to make any promises or threats for actions to be taken outside of the game. In the third game, rice payoffs were uncertain, dependent on a random weather shock. Payoffs were stochastic for both the clean and dirty canals, but the variance was higher when canals were dirty. The distribution of the payoffs was symmetric in both cases. The expected difference between the clean-canal payoff and the blocked-canal payoff was the same as in the non-stochastic game, so the expected returns to canal maintenance were held constant. This game was included to represent the importance of clean canals in reducing the negative impacts of weather shocks by reducing flood risk in order to improve the salience of the game as a learning tool.

At the end of the session, one round from each of the three games was chosen at random to determine final payoffs. Total payments ranged between approximately 150% and 250% of the typical daily agricultural wage in the region. All decisions and payments were kept private.

4 Data and Summary Statistics

4.1 Data sources

I draw data from four sources:

1. The baseline household survey of all study participants, conducted in February-March 2014. Study participants were drawn randomly from all farmers being included in the technology adoption project.

2. Data from the public goods experiments for all compliers in the treatment group, that is, those who were invited to the treatment and participated in the experiments.

3. Canal-cleaning data that includes the names of all farmers who participated in *kombits*, and the dates on which they participated.

4. A follow-up survey conducted several months after the experiments with the majority of study participants.

4.2 Summary of Household Baseline Characteristics

A balance test between the treatment and control group shows that the groups are well balanced on most household characteristics, providing support for the integrity of the randomization process (see Table 1). The dummy variables for small commerce, household business, and salaried labor refer to whether anyone in the household participates in that economic activity (households may participate in more than one). Because small commerce and small households business income includes expenses, some households report negative non-farm incomes as well as the more common negative farm incomes. Because many farmers are non-landowners, I report both land owned and land cultivated, as well as the share of cultivated land that a farmer owns. About half of farmers cultivate multiple parcels and in some cases these parcels are outside of our study area, so I report both the total land owned and cultivated and the land owned and cultivated in the study area.

The one variable that is statistically different between treatment and control groups is the number of parcels cultivated in the study area. This is due to the sampling design, in which I randomly selected 60 farmers and chose their closest neighbors to construct geographic clusters. Farmers with more than one parcel were more likely to be selected as neighbors, so I control for this variable in all analyses.

We need to be concerned about selection bias if, among those assigned treatment, compliers and non-compliers are systematically different. The two groups were found to be balanced on most observables, with several important differences (see Table 3). Several variables related to land cultivation and ownership are significantly different between the two groups, with compliers cultivating more parcels than non-compliers and non-compliers owning more land. I discuss later possible bias from unobservables that could come from farmers' self-selection into the treatment and correct for such a possibility using an Intention to Treat measure and Treatment on the Treated methods.

Table 1: Balance test on baseline characteristics between treatment and control groups

	Treatment	Control	Difference
Household size	5.02	4.74	0.28
	(2.50)	(2.48)	(0.19)
Age of HH head	53.5	53.5	0.026
	(12.4)	(13.7)	(1.02)
Sex of HH head	0.76	0.71	0.051
	(0.43)	(0.46)	(0.034)
Education of HH head	5.39	5.49	-0.098
	(4.39)	(4.46)	(0.34)
Farm profit (1,000 HTG)	59.7	43.8	15.9
	(223.7)	(264.0)	(19.5)
Nonfarm income (1,000 HTG)	102.7	82.7	19.9
	(485.9)	(509.7)	(38.6)
Total HH income (1,000 HTG)	196.1	171.3	24.9
	(557.9)	(784.5)	(56.0)
Small commerce (0/1)	0.45	0.44	0.0083
	(0.50)	(0.50)	(0.038)
Household business (0/1)	0.15	0.19	-0.042
	(0.36)	(0.40)	(0.030)
Salaried labor (0/1)	0.14	0.12	0.015
	(0.34)	(0.33)	(0.025)
Daily wage labor (0/1)	0.20	0.18	0.017
	(0.40)	(0.38)	(0.030)
Total land owned (1/100 ha)	27.4	29.4	-2.04
	(55.7)	(64.0)	(4.73)
Share of cultivated land owned	0.29	0.33	-0.036
	(0.41)	(0.43)	(0.032)
Total land cultivated (1/100 ha)	61.7	56.2	5.49
	(64.1)	(73.0)	(5.41)
Total parcels cultivated	2.45	2.32	0.13
	(1.52)	(1.54)	(0.12)
Land cultivated in study (1/100 ha)	49.8	43.6	6.13
	(54.1)	(49.1)	(3.89)
Parcels cultivated in study	2.01	1.82	0.19**
	(1.34)	(1.17)	(0.094)
N	561	243	804

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

This table provides sample statistics on the farmers' baseline demographic characteristics and differences, among all farmers selected for treatment, between compliers and non-compliers.

Table 3: Balance test on baseline characteristics between compliers and non-compliers

	Compliers	Non-Compliers	Difference
Household size	5.09	4.63	0.46
	(2.51)	(2.45)	(0.44)
Age of HH head	53.5	53.8	-0.27
	(12.0)	(14.8)	(2.19)
Sex of HH head	0.76	0.74	0.024
	(0.43)	(0.45)	(0.076)
Education of HH head	5.37	5.49	-0.11
	(4.42)	(4.29)	(0.79)
Farm profit (1,000 HTG)	64.4	33.7	30.7
	(241.3)	(70.2)	(40.0)
Nonfarm income (1,000 HTG)	113.8	42.6	71.2
	(525.0)	(141.6)	(85.9)
Total HH income (1,000 HTG)	213.6	100.9	112.7
	(600.8)	(177.4)	(99.7)
Small commerce (0/1)	0.46	0.37	0.095
	(0.50)	(0.49)	(0.088)
Household business (0/1)	0.14	0.21	-0.069
	(0.35)	(0.41)	(0.064)
Salaried labor (0/1)	0.15	0.053	0.099
	(0.36)	(0.23)	(0.060)
Daily wage labor (0/1)	0.21	0.11	0.11
	(0.41)	(0.31)	(0.070)
Total land owned (1/100 ha)	24.2	44.6	-20.4**
	(43.2)	(98.0)	(9.76)
Share of cultivated land owned	0.28	0.37	-0.089
	(0.39)	(0.48)	(0.072)
Total land cultivated (1/100 ha)	62.6	56.6	6.00
	(60.4)	(81.7)	(11.3)
Total parcels cultivated	2.55	1.92	0.63**
	(1.58)	(1.00)	(0.27)
Land cultivated in study (1/100 ha)	50.6	45.4	5.24
	(53.8)	(56.1)	(9.56)
Parcels cultivated in study	2.09	1.55	0.54**
	(1.41)	(0.76)	(0.23)
N	205	38	243

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

This table provides sample statistics on the farmers' baseline demographic characteristics and differences, among all farmers selected for treatment, between compliers and non-compliers.

4.3 Summary of Experiment Data

Overall cooperation levels were high in all rounds of the public goods games, and increased as the game was repeated. The game was designed to make coordination challenging, by setting the threshold for reaching the higher rice profit at 25, a number not divisible by four. The efficient average contribution, i.e. the minimum contribution necessary to reach the high-yield threshold, was 6.25, which some groups achieved in a "fair" manner by each contributing six and taking turns over the rounds giving seven. Average contributions were slightly below the efficient level in the first round, at 6.19 days, but in all subsequent rounds the average was above 6.25. Mean contributions over the seven rounds of the second game were 6.49, and over the seven rounds of the third game, 6.70. Average contributions rose slightly over the 15 rounds (see Figure 1), but the increase is not statistically significant. The variance in contributions did not change over time.

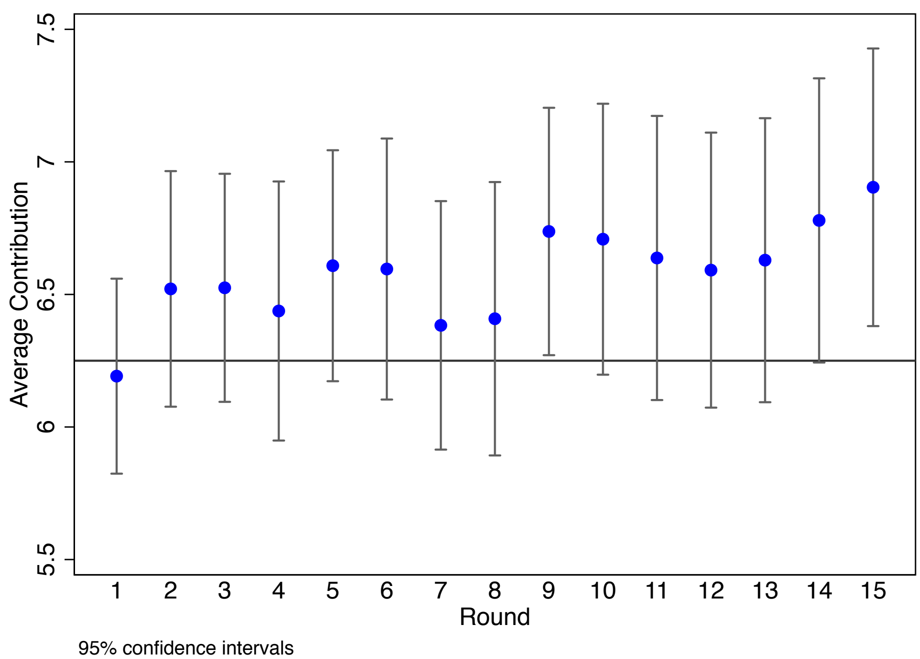

Figure 1: Average Individual Contributions to the Public Good
This figure shows average contributions, with 95% confidence intervals, for each round. Round 1 was the one-shot game with no communication; rounds 2-8 were the repeated game with verbal communication allowed between each round; rounds 9-15 were the repeated game with verbal communication and stochastic public goods payoffs.

As public goods contributions increased over rounds, average profits tended to increase as well as the game was repeated. These findings are consistent with previous literature demonstrating learning over the course of coordination games. I also find that the correlation between a participant's contribution to the public good and the contributions of others in the group increases over time, which suggests improved coordination among group members over repetitions of the game. These findings suggest that farmers may be learning over time the role the public good plays in their farm profits within the context of the game and learning more about how the other members of their group participate in provision of the public good.

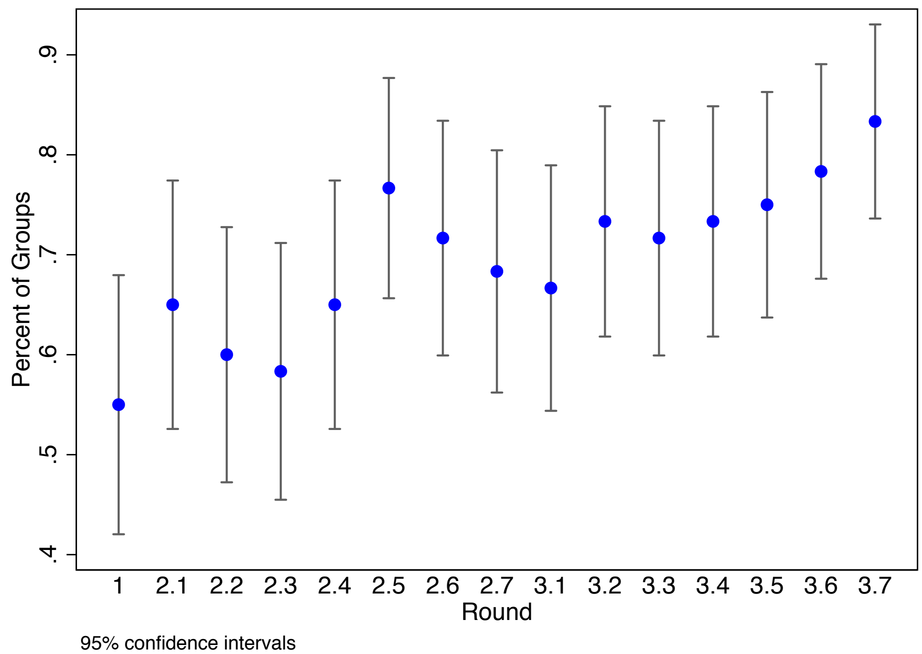

Figure 2: Percent of Groups Achieving the Cooperative Threshold
This figure shows the percent of groups that achieved the cooperative threshold, defined as the level of contributions to the public good that results in high payoffs from the public good for all participants. Round 1 was the one-shot game with no communication; rounds 2-8 were the repeated game with verbal communication allowed between each round; rounds 9-15 were the repeated game with verbal communication and stochastic public goods payoffs.

A more informative measure of successful cooperation than average contributions is the number of groups that cross the threshold between the low and high payoff, which I'll refer to as the *cooperative threshold* (see Figure 2). Here we also see improvement over time. In the first round, 55% of groups successfully achieve or exceed the cooperative contribution level. This number generally increased over the rounds, and in the final round 83% of groups cross the threshold. In this case the number of groups achieving cooperation in each of the last four rounds was statistically different from the first round.

4.4 Summary of Canal Cleaning Data

Of all of the households in the study, 11.5% participated in at least one day of canal cleaning during the planting season. Table 5 shows details on *kombit* participation for treated and control farmers, including both a measure for whether or not the farmer participated and measures of their level of participation. Not every farmer in our sample had the same opportunity to participate in the *kombits*, as each irrigation bloc organized its own cleaning schedule with a varying number of total days. This variation is due to differences in organizational capacity between the irrigation blocs as well as possible differences in the need for canal cleaning because of differences in drainage capacity and other physical characteristics. Due to this variation, another way to quantify an individual's level of participation that accounts for differences in opportunity to participate is to calculate the share of canal cleaning days organized in one's bloc in which the individual participated.

Table 5: Kombit participation among treated and control farmers

	Treatment	Control	Difference
Kombit (0/1)	0.14	0.093	0.051**
	(0.35)	(0.29)	(0.024)
Kombit Days	1.51	0.75	0.76***
	(3.72)	(2.48)	(0.22)
Kombit Share	0.12	0.063	0.053***
	(0.29)	(0.21)	(0.018)

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

This table shows participation in canal cleaning *kombits* for treated and control farmers. Kombit (0/1) is a dummy variable indicating whether a farmer participated in any canal cleaning. Kombit Days is the number of days farmers participated in canal cleaning. Kombit Share refers to the number of days that farmers participated, as a share of the number of days that were organized in each farmer's irrigation block.

5 Empirical Results

5.1 The Impact of Public Goods Games on Canal Cleaning

The primary question about the impact of exposure to public goods games on real-world public goods contributions is simply whether those exposed to the games were more likely to participate in any *kombits*, or canal cleaning. To investigate this question I estimate a linear probability model:

$$K_i = \beta_0 + \beta_1 T_i + X_i' \beta_2 + \varepsilon_i \tag{4}$$

where K_i is a dummy variable indicating whether farmer i participated in any *kombits* during the season, T_i indicates whether the farmer was exposed to the public goods games treatment, and X_i is a vector of household controls. The household controls included are: age of household head, sex of household head, household size, education of household head, farm profit, non-farm income, number of parcels cultivated (total and in the study), land cultivated (total and in the study), the share of parcels cultivated that are owned, and production block.

Because of the possible bias caused by non-compliance by some farmers assigned to the treatment group, I present both an Intention to Treat (ITT) analysis and a Treatment on the Treated (TOT) analysis. The compliance rate was approximately 80%. This rate is a result of the design, in which we invited five people for each group of four to ensure the appropriate number of people would participate. For the ITT specification, I define I_i to indicate whether farmer i was invited to participate in the PG games:

$$K_i = \beta_0 + \beta_1 I_i + X_i' \beta_2 + \varepsilon_i \tag{5}$$

In the TOT analysis, I define as treated only as those who received the treatment, i.e. those who were invited and participated in the games. Selection into participation among those invited causes concern about endogeneity, as possible systematic differences between those who agree to participate and those who refuse could bias our results. I therefore conduct the TOT analysis using an instrumental variables approach, with the invitation to participate as the exogenous instrument:

$$K_i = \beta_0 + \beta_1 \hat{P}_i + X_i' \beta_2 + \varepsilon_i \tag{6}$$

where $\hat{P}_i$ is the predicted probability of participation in the games based on the first-stage regression:[5]

$$P_i = \alpha_0 + \alpha_1 I_i + X_i' \alpha_2 + \epsilon_i \tag{7}$$

[5]The F statistics in the first-stage regression are 3476 and 182.5 with and without household controls, respectively.

Results of both approaches are presented in Table 7, with the second column for each specification displaying results from regression models that include household controls. I present ITT results in columns 1 and 2 and TOT results, using invitation to participate as an instrument, in columns 3 and 4. Because treatment was implemented to groups of four neighboring farmers, standard errors in all analyses are clustered at the experimental group level. Control farmers have been assigned to geographic clusters of the same size as those in the treatment group, based on the same maps that were used to create groups for the experiments. Such clustering allows us to correct for possible spatial correlation in errors, which could occur if local variation in water manageability or canal quality is correlated with the error.

Table 7: Probability of *Kombit* Participation

	ITT		TOT(IV)	
	Kombit (0/1)	Kombit (0/1)	Kombit (0/1)	Kombit (0/1)
PG Games: Invited	0.051*	0.048*		
	(0.028)	(0.028)		
PG Games			0.061*	0.057*
			(0.034)	(0.033)
Constant	0.093***	0.036	0.093***	0.039
	(0.012)	(0.064)	(0.012)	(0.063)
Household Controls	No	Yes	No	Yes
Observations	804	777	804	777
Adjusted R^2	0.005	0.042	0.008	0.046

Standard errors (in parentheses) are clustered at the group level

* p<.1, ** p<.05, *** p<.01

This table estimates the probability of participating in any canal cleaning *kombits* as a function of the experimental treatment. Household controls include age of household head, sex of household head, household size, education of household head, farm profit, non-farm income, number of parcels cultivated (total and in the study), land cultivated (total and in the study), the share of parcels cultivated that are owned, and production block.

The results show a weakly significant but substantial increase in the probability of participating in canal cleaning for farmers who participated in the PG games. Exposure to the games increased the likelihood of *kombit* participation by five to six percentage points, a substantial increase given the average participation rate of 11.5%. The ITT results are approximately 20% smaller in magnitude than the TOT results, as expected given the take-up rate, with a similar significance level.

5.2 Intensity of *Kombit* Participation

To examine the treatment effect in more detail, we can ask whether exposure to the PG games affects farmers' level of participation in the *kombits*. In this analysis, I focus on two outcome variables that measure the intensive margin of *kombit* participation. The first is the number of days that a farmer spent cleaning canals, which ranges among participants from 5 to 16. The second measure I use is the number of days spent cleaning as a share of the number of days organized in one's block, which accounts for the fact that farmers in different irrigation blocks were requested to participate for a different number of days.

With both of these measures, I use a Tobit model, where we can think of our desired outcome variable K_i^* as farmer i's propensity to participate in the *kombits*:

$$K_i^* = \beta_0 + \beta_1 T_i + X_i' \beta_2 + \varepsilon_i \tag{8}$$

but we only observe K_i, his actual participation level.

$$K_i = \begin{cases} K_i^* & \text{if } K_i^* > 0 \\ 0 & \text{if } K_i^* \leq 0 \end{cases} \tag{9}$$

The Tobit model accounts for the fact that all observations of zero are not equivalent. If we believe that those who participate in canal cleaning are somehow different from those who do not participate, a simple linear regression will not capture the distinction between participants and non-participants. Conceptually, we can think of a linear regression in this case as treating the distance between zero days of participation and one day the same as the distance between five and six days of participation. On the contrary, the move from zero days to one day is a move from non-participant to participant, so we need a model that treats positive values as fundamentally different from zero values. If we imagine that some non-participants are closer to participating than others, a small nudge toward participation may push the former group into participation while a bigger nudge would be required for the latter group. Mathematically, we can think of individuals with a $E(K_i^*|T_i = 0)$ negative but close to zero being only a short distance from becoming contributors, so they will be more likely to respond to treatment by become contributors than individuals with $E(K_i^*|T_i = 0)$ large and negative.[6]

[6] I choose a Tobit model rather than a Heckman-style selection model that specifies two distinct processes for the selection equation and the outcome equation because in this context, there are no variables that can reasonably be excluded from the outcome equation, as required by a selection model. In some settings there are factors that influence the decision to participate but not the decision regarding how much to participate. The decision to participate in the

We cannot model the effect of the treatment on participation levels by looking only at those who participate, because selection into participation in the canal cleaning days is not random, as I showed in the first result. The treatment has a clear impact on whether or not one chooses to participate in canal cleaning, and is likely to induce what we may call "marginal" participators into canal cleaning; these are farmers who would not have participated had they not been treated with the PG games. We would expect participation levels among these marginal participators to be lower than among always-participators, i.e. those who would have participated regardless of treatment status, so we would expect the estimated treatment effect on participation levels among participants only, without accounting for the distinction between participants and non-participants, to be biased downward.

In this section, I conduct only an ITT analysis because instrumental variables analysis is biased in a Tobit model with a binary endogenous regressor. The ITT measure is more conservative than the TOT measure, so we could consider the results to be a lower bound for the effect. The treatment has a large impact on the amount of time spent participating in *kombits*: approximately four days, or in the share model, one-third of the number of days organized in one's block. The results are presented in Table 8.

Table 8: Level of Participation in Canal Cleaning: Tobit Model (ITT)

	Kombit Days	Kombit Days	Kombit Share	Kombit Share
PG Games: Invited	4.82***	4.09**	0.38**	0.31**
	(1.72)	(1.59)	(0.15)	(0.14)
Constant	-18.6***	-22.1***	-1.59***	-2.07***
	(1.62)	(4.90)	(0.14)	(0.43)
Household Controls	No	Yes	No	Yes
Observations	804	777	804	777

Standard errors (in parentheses) are clustered at the group level

* p<.1, ** p<.05, *** p<.01

This table estimates the level of participation in canal cleaning *kombits* as a function of the experimental treatment. Columns 1 and 2 use as the outcome variable the number of days spent canal cleaning, while columns 3 and 4 use the number of days as a share of the total number of days organized in each farmer's canal. Household controls include age of household head, sex of household head, household size, education of household head, farm profit, non-farm income, number of parcels cultivated (total and in the study), land cultivated (total and in the study), the share of parcels cultivated that are owned, and production block.

It is clear that, regardless of which outcome measure we use, exposure to the public goods games

job market, for example, may be influenced by fixed costs to searching for a job that do not affect the hours worked once one has joined the job market. In this setting, no such exclusion restrictions seem valid. For a robustness check, I conducted the same analysis using the Heckman model and the results were qualitatively similar.

leads to a substantial increase in *kombit* participation. What these results don't tell us, however, is *why* we observe the results we do. As discussed in Section 2, there are several possible pathways through which exposure to the collective action dilemma in the form of a field experiment may lead to behavior change. In the next section, I examine possible mechanisms that could be driving the treatment effect.

6 Mechanisms

We can use differences observed within the experiments as well as data from other surveys conducted with our sample to explore the reasons why we see the impact on *kombit* participation that we do. The set-up of the game, framed to approximate the real relationship between canal cleaning and rice production, may teach farmers something about the technical, physical relationships between rice yields, canal functionality, and labor inputs into canal maintenance. Farmers may learn about the role of canal performance in the production function for rice, $Y_i = f(x_i, g(Z))$, or about the role of time spent canal cleaning in the production function for canal performance, $g(Z)$. If this is the case, we can see the experimental treatment as a pedagogical tool to improve farmers' knowledge of the relevant production functions.

We can also think about the games as an opportunity to learn about a *social technology*. A key input into the functionality of a canal is the labor contributed by other individuals with whom one shares the canals. An individual's expectations of yields depends on his understanding of the technical production function as well as his expectations of the behavior of the others with whom he shares a canal. The experiments may affect the latter if they allow him to learn something about his plot neighbors. Finally, the utility that farmers obtain from contributing to a public good may depend on characteristics of or his relationships with those with whom he shares the public good. The experiments provide a venue to learn more about those who benefit from his cooperative behavior, which could, in itself, influence his contributions.

In this section I explore technical and social mechanisms to provide some suggestive evidence of the mechanism driving the treatment effect.

6.1 Technical Learning About the Production Functions

One way the experimental treatment could influence behavior is by acting as a pedagogical tool to improve understanding of the trade-offs between private work and contributions to the public good and the benefits of the public good: the yield production function, $Y_i = f(x_i, g(Z))$ or the production function of the public good itself $g(Z)$. If participation in the games teaches treated farmers about the benefits of clean canals, this increased awareness should increase contributions to the public good.

Optimization of inputs into agriculture depends on knowledge of the production function, which is not known ex ante. There is a growing literature on how farmers learn about production functions, both from their own experience, (e.g. Foster and Rosenzweig (1995)) and from the experience of others (e.g. Besley and Case (1993); Conley and Udry (2001)). These learning models typically assume that data, either from one's own farm or neighbors' farms, is they key input into learning. However, examples of individuals failing to assimilate all of the available information into their understanding of production functions suggests that the availability of data is not sufficient for learning. In a model of incomplete learning, Hanna, Mullainathan, and Schwartzstein (2014) show that farmers selectively choose which information to pay attention to when their production function depends on many inputs. Failing to notice a key input leads farmers to produce far below their productivity frontier despite the availability of data on the ignored input. In a similar manner, farmers in our study may, despite their experience and the availability of information on the role of water management in rice production, ignore its importance as an input when making their time allocation decisions. The experiments, by highlighting the importance of water management, may prime farmers to pay more attention to this input.

To examine whether the experiments served to increase understanding of the role of management in rice production, we asked farmers to estimate the production functions in order to compare responses between treatment and control farmers. In a follow-up survey conducted several months after the PG games were conducted, we asked farmers to estimate how their yields would change as a function of the performance of the canal running closest to their parcels. Farmers were able to estimate these differences without much trouble: only a handful of farmers responded with "I don't know" and only one estimated higher yields when the canals were blocked than when they were clean. Figure 3 shows density curves for expected yields at each level of canal functionality, reported as the share of the maximum yield farmers expected when canals are at the highest level. Pairwise t-tests between means show that the mean for each level of canal performance is statistically different from each other level, showing some consistency among farmers in estimating the relationship between canal performance and expected yields.

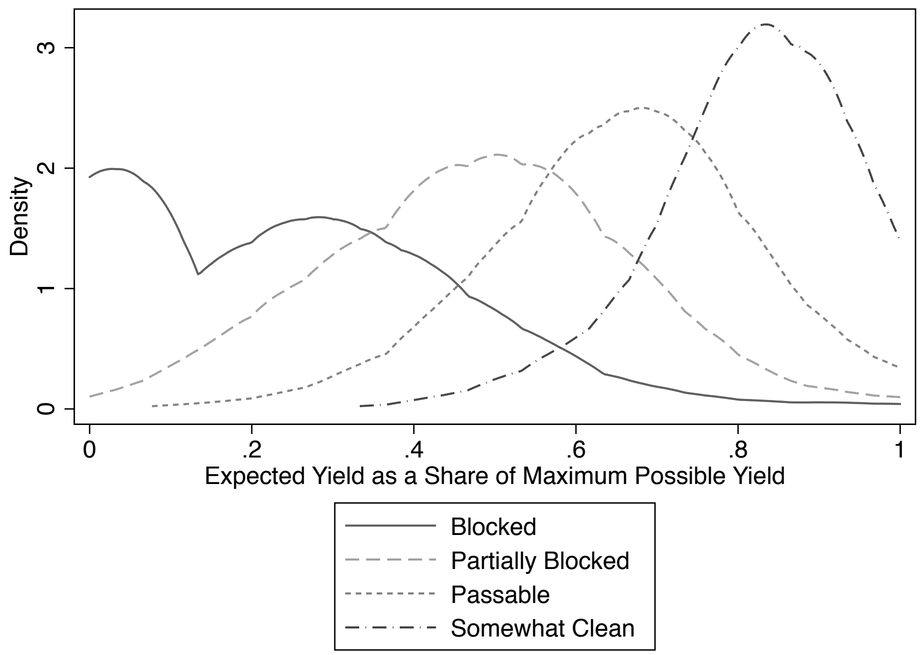

Figure 3: Yield production function estimates
Farmers estimated their expected yields on their main field for five levels of canal performance (blocked, partially blocked, passable, somewhat clean, and clean). This figure represents estimates for each level as a share of the expected yield when canals are at the cleanest level.

If we hypothesize that the PG games changed *kombit* participation by changing something about farmers' understanding of the yield production function, we would expect to see differences in how canal performance affects expected yields between treatment and control farmers. I test whether this is the case with a simple model:

$$EY_i = \beta_0 + \beta_1 C + \beta_2 T_i + \beta_3 T_i * C + X_i' \beta_4 + \varepsilon_i \tag{10}$$

where EY_i is the farmer's estimated yield, as a share of the maximum possible yield, C is the level of canal performance, and $T_i * C$ is an interaction term between treatment and the canal performance level. Note that in this specification, each farmer has four observations: one for each canal performance level other than the highest level.

If the treatment changed farmers' perceptions of the effect of canal performance on yields, we would expect a significant coefficient on the interaction term, β_3. Instead, we observe that while farmers attribute higher yields to better-performing canals (the coefficient β_1 is positive and strongly statistically significant), we do not observe any difference between the treatment and control group (see Table 9). These results do not provide any evidence for a treatment effect on learning about the relationship between canals and yields.

Table 9: Yield Estimates as a Function of Canal Performance: Testing for Treatment Effect

	Yield Estimate as a Share of Maximum	Yield Estimate as a Share of Maximum
Canal Cleanliness Level	0.19***	0.19***
	(0.0026)	(0.0026)
PG Games: Invited	-0.027	-0.027
	(0.023)	(0.023)
Canal Level * Treatment	0.0043	0.0038
	(0.0046)	(0.0048)
Constant	0.083***	0.084***
	(0.012)	(0.032)
Household Controls	No	Yes
Observations	2914	2814
Adjusted R^2	0.744	0.743

Standard errors (in parentheses) are clustered at the group level

* p<.1, ** p<.05, *** p<.01

Farmers estimated their expected yields on their main field for five levels of canal performance (blocked, partially blocked, passable, somewhat clean, and clean). This table estimates the farmers' expected yields for each level as a share of the expected yield when canals are at the cleanest level.

We can also look at farmers' estimation of the public goods production function, that is, the performance of the canals as a function of the amount of time spent cleaning them. In the same follow-up survey, we asked farmers to estimate how many days each farmer in the production block would have to spend cleaning to bring a canal to a functional level, depending on its current level. A density curve for each initial cleanliness level is shown in Figure 4. Again, a t-test between each level shows that the mean number of days farmers estimate are required for each initial cleanliness level is statistically different; in other words, the number of days required to clean a blocked canal is statistically different from the number of days required to clean a somewhat blocked canal, and so on.

As with the rice yield production function, if we think that the treatment changed farmer's understanding of the relationship between time spent cleaning the canals and their functionality,

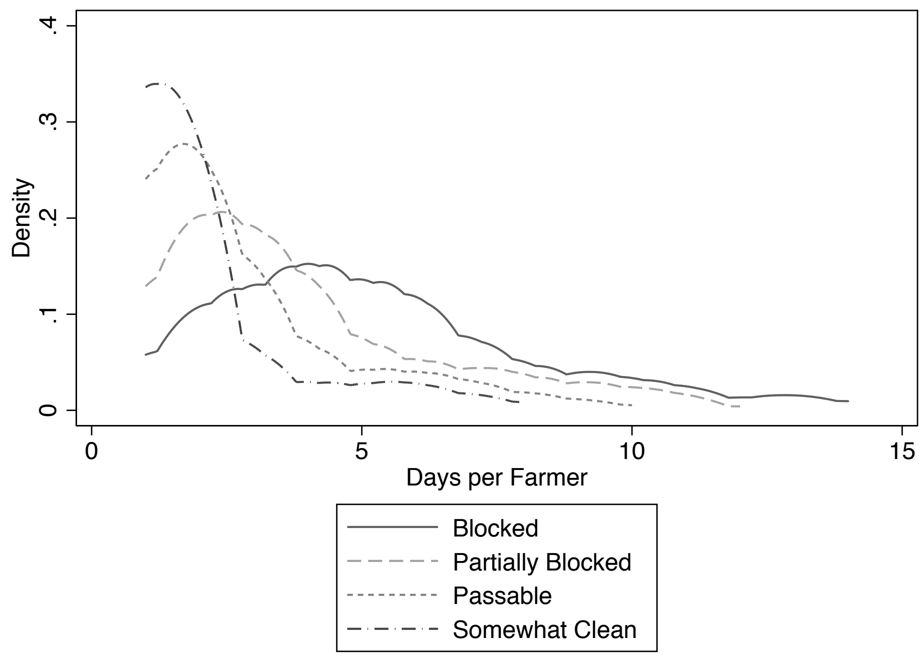

Figure 4: Public good production function estimates
Farmers estimated the labor require to clean canals from four different starting points (blocked, partially blocked, passable, and somewhat clean). This figure represents estimates for each level of canal cleanliness and tests whether the estimated relationship between yields and canal performance vary by treatment status.

we would expect to see a difference between the treatment and control groups in this question. We can test this with a similar specification to the test of the yield production function:

$$ED_i = \beta_0 + \beta_1 C + \beta_2 T_i + \beta_3 T_i * C + X_i' \beta_4 + \varepsilon_i \tag{11}$$

where ED_i is farmer i's estimate of the number of days of labor each person in the block would have to contribute to clean the canal, based on its initial level of cleanliness. If the PG games changed farmers' perceptions of the public goods production function, we would expect a significant coefficient on the interaction term. As with the yield production function, we observe a clear relationship between the days required to clean the canal and its initial cleanliness level, but we do not observe that this relationship depends on treatment status (see Table 10).

While these results do not provide conclusive evidence that learning is not taking place, I cannot reject the null hypothesis that farmers who received the experimental treatment have the same understanding of both production functions as those in the control group. This result suggests

Table 10: Canal Performance as a Function of Labor Input: Testing for Treatment Effect

	Labor Required to Clean Canal	Labor Required to Clean Canal
Canal Cleanliness Level	-1.41***	-1.41***
	(0.059)	(0.060)
PG Games: Invited	-0.15	-0.24
	(0.64)	(0.64)
Canal Level * Treatment	0.025	0.028
	(0.14)	(0.15)
Constant	7.14***	8.90***
	(0.26)	(0.88)
Household Controls	No	Yes
Observations	2321	2241
Adjusted R^2	0.192	0.220

Standard errors (in parentheses) are clustered at the group level

* p<.1, ** p<.05, *** p<.01

Farmers estimated the labor require to clean canals from four different starting points (blocked, partially blocked, passable, and somewhat clean). This table estimates the farmers' expected labor requirements for each level of canal cleanliness and tests whether the estimated labor requirements for canal cleaning vary by treatment status.

that learning may not be the primary mechanism driving the observed treatment effect.

6.2 Social Learning About Plot Neighbors

The link between social capital and collective action has been documented in many contexts, demonstrating that social bonds and norms can often be more effective tools for managing shared resources than formal rules and regulations (Sethi and Somanathan (1996); Pretty (2003); Ostrom (2014)). Social capital has been defined in a number of ways, but it typically includes elements of social networks and social norms that enable people to act collectively. In this section I explore these two elements to study whether the experimental treatment had any effect on either norms or social networks, and whether these characteristics affect behavior.

6.2.1 Social Norms

If conforming to a behavioral norm is an important driver of one's decision to participate in canal cleaning, an important piece of information that one may learn from participation in the games is what kind of behavior they can expect from their neighbors. Participants who learn that their neighbors are unlikely to contribute to communal canal-cleaning days may experience a different treatment effect than those who learn that their neighbors are likely to contribute.

To analyze the differential treatment effect by neighbors' contributions, I separate the treatment group into two groups, based on the initial contributions of other members of the group during the PG games:

1. High-contribution treatment: I define TH_i as a dummy variable indicating that the average contribution of all other members of farmer i's group (excluding i's own contribution) in the first round was above the optimal contribution. Recall that the optimal average contribution is 6.25.

2. Low-contribution treatment: I define TL_i as a dummy indicating that the average contribution of other members in the first round was below the optimum of 6.25

For farmers in the treatment group, either $TH_i = 1$ or $TL_i = 1$, while for farmers in the control group, $TH_i = TL_i = 0$. With these new treatment variables, my estimating equation is:

$$K_i = \beta_0 + \beta_1 TH_i + \beta_2 TL_i + X_i' \beta_3 + \varepsilon_i \tag{12}$$

In any analysis that includes behavior in the PG games as part of the treatment, we have to be concerned about endogeneity. Farmers who contribute more of their endowment to the public good during the games were found to contribute more time to canal cleaning in real life, but this is likely to be simply an indication that cooperative individuals cooperate more in both contexts.[7] In this model use the initial contributions of others in the group in this section as it is the only measure from the game that has not been influenced directly by the game behavior of other farmers in the group. After the first round, each participant's behavior may be a reaction to the behavior of others in the group, so all decisions made after the first round are at least partly endogenously determined. By considering only the initial contributions of others, we can divide the treatment group into two in order to get two separate treatment effects: the effect of playing the game with a group of high contributors and the effect of playing the game with a group of low contributors.

The challenge with using in-game behavior as part of the treatment is that we only have such data for compliers. As previously discussed, excluding those who were invited but did not participate could result in bias if compliers are systematically different from non-compliers. To solve this

[7] In a simple model estimating the probability of participation in *kombits* separately for treated farmers who contributed above the optimal contribution level on average, and those who contributed below on average, I find that the former group is 10-11 percentage points more likely to participate in *kombits*, while I find no treatment effect for the latter group.

problem, I define the treatment variable for non-compliers as the average of the treatment variable for members of the same group. Each person who was invited to participate was assigned to a group in which he would have participated, based on geographic clusters. I therefore can use the average contributions of that group to assign a treatment level even to those who did not participate, as the treatment they would have received had they complied. I can use this measure to include all invited farmers in the analysis, in both an ITT measure and an Instrumental Variables analysis using the invitation to participate multiplied by the hypothetical treatment level as an instrument.[8]

Table 11: Treatment Effect Varies Depending on Initial Contribution Level of Others in Group

	ITT		TOT(IV)	
	Kombit (0/1)	Kombit (0/1)	Kombit (0/1)	Kombit (0/1)
Others Above Optimum	0.091**	0.071*	0.11**	0.082*
	(0.041)	(0.040)	(0.048)	(0.045)
Others Below Optimum	0.014	0.026	0.020	0.032
	(0.035)	(0.036)	(0.039)	(0.041)
Constant	0.092***	0.030	0.092***	0.033
	(0.012)	(0.064)	(0.012)	(0.063)
Household Controls	No	Yes	No	Yes
Observations	809	781	809	781
Adjusted R^2	0.008	0.045	0.011	0.049

Standard errors (in parentheses) are clustered at the group level

* p<.1, ** p<.05, *** p<.01

In this table I re-estimate the primary model testing the treatment effect on the probability of participation in *kombits* with the treatment group separated into two groups: farmers in groups in which the average contributions of other group members in the first round was above the optimal contribution level, and those in groups in which the average contributions we below the optimum.

When we allow the treatment effect to vary by initial impressions of other group members' contributions, we see that the increase in real-world participation in *kombits* is present only for those whose group members are high contributors. In Table 11 I show the result with the probability of any *kombit* participation as the outcome variable. For those who were in groups where other group members contributed above the optimum, the increase in the probability of participating in *kombits* is higher in magnitude than the initial result with all treated farmers grouped together. For those in groups where others contributed below the optimum, we do not see a statistically significant effect.

[8]As robustness checks I've estimated the following alternative specifications: dropping all of the non-compliers from the analysis, running two separate regressions with the two separate treatment groups. Both alternatives yield the same results as the estimation presented here.

We observe a very similar result when we consider the level of *kombit* participation as the outcome variable: a strong, statistically significant, relationship between the high-contribution treatment and participation, and little evidence of a relationship between the low-contribution treatment and participation. I again use a Tobit ITT model to examine the effect on the number of days of participation (see Table 12).

Table 12: Treatment Effect Varies Depending on Initial Contribution Level of Others in Group (Level of Participation: Tobit-ITT)

	Kombit Days	Kombit Days	Kombit Share	Kombit Share
Others Above Optimum	6.68***	5.05***	0.53***	0.40**
	(2.04)	(1.90)	(0.18)	(0.17)
Others Below Optimum	2.85	2.97	0.22	0.22
	(2.47)	(2.29)	(0.21)	(0.20)
Constant	-18.6***	-22.4***	-1.59***	-2.10***
	(1.62)	(4.88)	(0.14)	(0.43)
Observations	809	781	809	781

Standard errors (in parentheses) are clustered at the group level

* p<.1, ** p<.05, *** p<.01

In this table I estimate the treatment effect on the level of participation in *kombits*, using both measures of participation level (days and share of the total days organized in one's block) with the treatment group separated into two groups: farmers in groups in which the average contributions of other group members in the first round was above the optimal contribution level, and those in groups in which the average contributions we below the optimum.

This finding suggests that expectations of others' participation in public goods provision may explain the mechanism driving the treatment effect we see. As discussed in Section 2, information about others' behavior could matter if conforming to a norm is important to individuals. If one places value on contributing at a similar level to one's neighbors, learning that others are likely to contribute at a high level is likely to increase contributions.

While this analysis uses only others' contributions in the first round because of the concern about exogeneity of later behavior within the game, first-round behavior is a strong predictor of later behavior throughout the game, as shown in Table 13. We can see here that the probability of the group arriving at or above the optimal contribution level by the end of the game is significantly and positively correlated with first-round behavior. I use two different measures of first-round behavior: the total contribution level in the first game (shown in Columns 1 and 2) and a dummy indicating whether or not the group contribution level was at or above the optimum in the first round (Columns 3 and 4). Both measures provide similar results.

We can examine the in-game data on how individuals respond to others' behavior as further

Table 13: Final Contributions as a Function of First-Round Contributions

Dependent Variable: Final Round Above Optimum				
Total Contributions: Game 1	0.024***	0.025***		
	(0.0062)	(0.0063)		
First Round Above Optimum			0.40***	0.42***
			(0.064)	(0.065)
Constant	0.100	0.25	0.41***	0.52***
	(0.19)	(0.27)	(0.087)	(0.19)
Household Controls	No	Yes	No	Yes
Observations	210	202	212	204
Adjusted R^2	0.071	0.197	0.161	0.282

Standard errors in parentheses

* p<.1, ** p<.05, *** p<.01

evidence of a propensity to conform to a common norm. Treating the within-game behavior as time series data, we can analyze how an individual's contributions to the public good evolve over the rounds of the game in response to the contributions of others with a simple specification:

$$PG_{it} = \phi_0 + \phi_1 PG_{it-1} + \phi_2 O_{it-1} + X_i' \phi_3 + \varepsilon_{it} \tag{13}$$

where PG_{it} is farmer i's contribution to the public good in time t and O_{it} is the average contribution of other members of farmer i's group in time t. The estimation results are presented in Table 14. We see clearly that both one's own behavior in previous rounds and the behavior of others are strong predictors of behavior in subsequent rounds.

Table 14: In-Game Contributions Over Time

	Contributions to PG	Contributions to PG
Contributions to PG *(t-1)*	0.57***	0.58***
	(0.036)	(0.035)
Avg PG Contribution of Group Members *(t-1)*	0.24***	0.22***
	(0.043)	(0.037)
Constant	1.28***	1.24***
	(0.31)	(0.43)
Household Controls	No	Yes
Observations	3346	2758
Adjusted R^2	0.471	0.485

Standard errors in parentheses

* p<.1, ** p<.05, *** p<.01

This finding supports the assertion that the behavior of others is a driver of one's own behavior, which may explain why those who learn that the farmers with whom they share a canal are more cooperative within the game are in turn morel likely to contribute at higher levels after the games are over.

6.2.2 Social Networks

A related potential social effect of the treatment is the social connectedness between farmers. The precise definition of social capital is heavily debated, but it typically includes elements of trust or norms as well as social networks. To explore this latter element, we surveyed farmers several months after the PG games to measure their level of connectedness to their plot neighbors.

Each farmer was asked four questions about each other person in their geographic cluster: whether the person is a plot neighbor, whether they ask the person for or give advice on farming matters,[9] and whether they share labor. Note that all of the people about whom these questions were asked were plot neighbors according to our maps of the study area, so the first question is a test of whether the farmer knows the other person farms a plot near them. It is possible for a farmer to answer yes to another social network question without knowing whether the farmer in question was a plot neighbor: 31% of farmers answered yes to at least one other question about another farmer they knew without identifying that person as a plot neighbor. In Table 15 I show a summary of responses to the social networks questions for treated and control farmers with a pairwise test of means between the two groups.

Table 15: Social Networks: Summary and T-Test

	Treatment	Control	Difference
Knows Plot Neighbors	0.63	0.49	0.14***
	(0.34)	(0.35)	(0.030)
Asks/Gives Farm Advice	0.56	0.42	0.14***
	(0.37)	(0.36)	(0.032)
Shares Labor	0.21	0.15	0.063**
	(0.34)	(0.28)	(0.026)
Any Soc. Net. Question	0.68	0.54	0.13***
	(0.33)	(0.35)	(0.030)

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

[9] There was almost complete overlap between these two responses, so I have grouped them into a single variable.

I define a social network variable for each question as the share of the other people in one's group for whom one answers yes to the question. I construct the variable as a share because the group size varied slightly. Most groups were five people, so each person was asked the social network questions about four others; if an individual identified three of the other four as plot neighbors, the variable "Knows Plot Neighbors" is equal to 0.75. I additionally define a variable "Any Social Network Question" as the share of people in one's group for whom the farmer answered yes to at least one of the four questions.

To test the effect of the treatment on social connectedness, I run separate regressions of each social network variable on the treatment. The results are presented in Table 17. Here I show the ITT analysis, but the result is nearly identical for TOT analysis using the invitation to participate as the instrument for participation. The effect of the treatment on social networks is quite strong: participation in the games increases the share of people an individual identifies for each measure by around 50%.

Table 17: Treatment Effect on Social Networks

	Knows Plot Neighbors	Asks/Gives Farm Advice	Shares Labor	Any Soc. Net. Question
PG Games: Invited	0.15***	0.13***	0.071**	0.13***
	(0.034)	(0.035)	(0.029)	(0.034)
Constant	0.32***	0.28***	0.18**	0.42***
	(0.083)	(0.089)	(0.071)	(0.085)
Household Controls	Yes	Yes	Yes	Yes
Observations	584	584	584	584
Adjusted R^2	0.070	0.050	0.013	0.043

Standard errors (in parentheses) clustered at the group level
* p<.1, ** p<.05, *** p<.01

As a robustness check, we can look at social connectedness of treated and control farmers before the experiments using similar social network questions asked at baseline. While the questions were similar, it is important to note that they were asked of five randomly selected farmers in the production block, not specifically of parcel neighbors. Nonetheless, this check provides evidence that the treatment was uncorrelated with initial levels of social connectedness in general, so the higher responses to social network questions among treatment farmers appear to be caused by exposure to neighbors in the PG games (see Table 18).

Of course, the effect of the treatment on social connectedness only matters if connectedness itself is actually related to the variable of interest, participation in *kombits*. In Table 20 I show that this is the case. All of the social network variables are correlated with greater participation in

Table 18: Social Networks Robustness Check: Summary and T-Test

	Treatment	Control	Difference
Related	0.024	0.025	-0.0011
	(0.12)	(0.12)	(0.0093)
Asks Farm Advice	0.033	0.026	0.0065
	(0.16)	(0.14)	(0.011)
Gives Farm Advice	0.033	0.030	0.0030
	(0.15)	(0.15)	(0.012)
Shares Labor	0.020	0.019	0.00050
	(0.12)	(0.11)	(0.0089)
Any Soc. Net. Question	0.086	0.076	0.010
	(0.42)	(0.38)	(0.030)

Standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

canal-cleaning days. To interpret the magnitude of the coefficients, we can think about the typical farmer who has four other farmers in his group. An increase of one person with whom the farmer says he shares communal labor, i.e. a 0.25 increase in the share, is associated with an increase in the probability of *kombit* participation of four percentage points. This is a substantial increase given the average participation rate of 11.5%, implying that one new connection with a neighbor is correlated with a 35% increase in *kombit* participation.

Table 20: Participation Effect of Social Networks

	Kombit (0/1)	Kombit (0/1)	Kombit (0/1)	Kombit (0/1)
Knows Plot Neighbors	0.074*			
	(0.041)			
Asks/Gives Farm Advice		0.14***		
		(0.037)		
Shares Labor			0.16***	
			(0.057)	
Any Soc. Net. Question				0.13***
				(0.041)
Constant	-0.018	-0.034	-0.023	-0.049
	(0.079)	(0.076)	(0.078)	(0.078)
Household Controls	Yes	Yes	Yes	Yes
Observations	584	584	584	584
Adjusted R^2	0.063	0.078	0.076	0.074

Standard errors (in parentheses) clustered at the group level
* p<.1, ** p<.05, *** p<.01

While I fail to find convincing evidence supporting a technical mechanism driving the treatment effect, I do find support for social mechanisms operating both through norms and social networks. It is important to note that these two mechanisms are not mutually exclusive; farmers may be learning both about the technical relationship between canal function and yields and about their neighbors at the same time. And though I fail to reject the null hypothesis that there is no impact on learning, we cannot rule out that learning may, in addition to the effect on social norms and social networks, be playing a role in the treatment effect we see.

7 Conclusion

This paper examines whether experimental lab-in-the-field games can serve as a pedagogical tool to change real-world behavior in a context similar to that modeled in the games. While previous studies have connected in-experiment behavior to real-world decisions, using experiments to influence decision-making outside of the game is a new area of research that could be promising for policy interventions aimed at changing similar behaviors.

Voluntary contributions to a public good provide just one example of a setting where behavior often does not conform to neoclassical predictions: we very often see, in experiments and in the real world, that people contribute and cooperate at a higher rate than one might predict based only on the expected financial returns to cooperation. Such behavior is highly variable across different settings, so behavioral interventions aimed at increasing public goods contributions have the possibility to improve welfare substantially in settings where contributions are sub-optimal.

In this paper I find that exposure to a public goods games designed to replicate the strategic considerations farmers face in the real world increases the average probability of contributing to a public good by approximately 66%. In terms of the total amount of labor each farmer contributed to canal cleaning, the treatment doubled the average contribution. Measuring a direct increase on canal performance would require randomization at the canal level, but this finding demonstrates a first step in the effectiveness of experimental games as a way to shift behavior, at least in the short run. To put the increase into context, farmers reported in a follow-up survey that the observed increase in average contributions would improve the performance of the canals enough to increase average yields on their plots by more than 10%.

By examining in-game behavior and survey data, I provide suggestive evidence supporting the hypothesis that the treatment effect is operating through social learning in the games. Farmers assigned to treatment groups with others who contribute at a high level during the experiments increased their real-world contributions, while we do not observe a treatment effect for those assigned to groups with low in-game contributors. This finding suggests that farmers are learning something about the other farmers with whom they share canals that determines how they will interpret the lessons of the game: those who learn that others are likely to contribute in turn contribute more themselves. Additionally, we see an effect on social networks among treated farmers. Participation in the games increased connectedness among farmers who cultivate plots near one another, and social connectedness to neighboring farmers is a strong predictor of contributions to the public good.

The evidence I provide in this paper suggests that experimental games could be a valuable tool to be used in development interventions to increase collective action around public goods provision. Conducting public goods games can be very inexpensive if run by local partners for the purpose of learning, as opposed to rigorous data collection. In fact, during the course of this research, the local irrigation association asked to be trained in running the games in subsequent years because he saw their possible value as a way to motivate farmers to participate in canal cleaning. More research is needed to explore the longer term impacts of such an intervention and to explore more rigorously the mechanisms behind the treatment effect to help us understand better why such an intervention works and how long we can expect the effects to last.

References

Allcott, H., and T. Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review* 104:3003–37.

Andreoni, J. 1990. "Impure altruism and donations to public goods: a theory of warm-glow giving." *The economic journal*, pp. 464–477.

Avdeenko, A., and M.J. Gilligan. 2014. "International interventions to build social capital: evidence from a field experiment in Sudan." Working paper No. 6772, World Bank Policy Research Working Paper.

Baland, J.M., and J.P. Platteau. 1998. "Wealth inequality and efficiency in the Commons, part II: the regulated case." *Oxford Economic Papers* 50:1–22.

Barr, A. 2003. "Trust and expected trustworthiness: experimental evidence from zimbabwean villages*." *The Economic Journal* 113:614–630.

Beath, A., F. Christia, and R. Enikolopov. 2013. "Do elected councils improve governance? experimental evidence on local institutions in Afghanistan.", pp. .

Besley, T., and A. Case. 1993. "Modeling technology adoption in developing countries." *The American Economic Review*, pp. 396–402.

Bó, P.D., and G.R. Fréchette. 2011. "The evolution of cooperation in infinitely repeated games: Experimental evidence." *The American Economic Review* 101:411–429.

Cardenas, J.C., and J. Carpenter. 2005. "Three themes on field experiments and economic development." *Field experiments in economics, JAI Press, Greenwich*, pp. 71–124.

Cárdenas, J.C., L.A. Rodriguez, and N. Johnson. 2011. "Collective action for watershed management: field experiments in Colombia and Kenya." *Environment and Development Economics* 16:275–303.

Carter, M.R., and M. Castillo. 2011. "Trustworthiness and social capital in South Africa: Analysis of actual living standards data and artifactual field experiments." *Economic Development and Cultural Change* 59:695–722.

Casey, K., R. Glennerster, and E. Miguel. 2012. "Reshaping institutions: Evidence on aid impacts using a preanalysis plan." *The Quarterly Journal of Economics* 1755:1812.

Chaudhuri, A. 2011. "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature." *Experimental Economics* 14:47–83.

Conley, T., and C. Udry. 2001. "Social learning through networks: The adoption of new agricultural technologies in Ghana." *American Journal of Agricultural Economics* 83:668–673.

Duffy, J., and J. Ochs. 2009. "Cooperative behavior and the frequency of social interaction." *Games and Economic Behavior* 66:785–812.

Fearon, J.D., M. Humphreys, and J.M. Weinstein. 2009. "Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict Liberia." *The American Economic Review*, pp. 287–291.

Ferraro, P.J., and M.K. Price. 2013. "Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment." *Review of Economics and Statistics* 95:64–73.

Fischbacher, U., S. Gächter, and E. Fehr. 2001. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics Letters* 71:397–404.

Foster, A.D., and M.R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103:pp. 1176–1209.

Hanna, R., S. Mullainathan, and J. Schwartzstein. 2014. "Learning through noticing: Theory and evidence from a field experiment." *The Quarterly Journal of Economics* 129:1311–1353.

Hardin, G. 1968. "The tragedy of the commons." *science* 162:1243–1248.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath. 2001. "In search of homo economicus: behavioral experiments in 15 small-scale societies." *American Economic Review*, pp. 73–78.

Hess, C. 1999. *A comprehensive bibliogcraphy of common pool resources*. Bloomington: Indiana University, Workshop in Political Theory and Policy Analysis.

Hoff, K., and P. Pandey. 2006. "Discrimination, social identity, and durable inequalities." *The American economic review*, pp. 206–211.

Isaac, R.M., J.M. Walker, and A.W. Williams. 1994. "Group size and the voluntary provision of public goods: experimental evidence utilizing large groups." *Journal of public Economics* 54:1–36.

Karlan, D.S. 2005. "Using experimental economics to measure social capital and predict financial decisions." *American Economic Review*, pp. 1688–1699.

Ledyard, J.O. 1995. "Public Goods: A Survey of Experimental Research. S. 111–194 in: JH Kagel/AE Roth (Hrsg.), The Handbook of Experimental Economics."

Mansuri, G., and V. Rao. 2012. *Localizing development: does participation work?*. World Bank Publications.

McCarthy, N., E. Sadoulet, and A. de Janvry. 2001. "Common Pool Resource Appropriation under Costly Cooperation." *Journal of Environmental Economics and Management* 42(3):297 – 309.

Moser, C.M., and C.B. Barrett. 2003. "The disappointing adoption dynamics of a yield-increasing, low external-input technology: the case of SRI in Madagascar." *Agricultural Systems* 76:1085 – 1100.

Ostrom, E. 2014. "Collective action and the evolution of social norms." *Journal of Natural Resources Policy Research* 6:235–252.

—. 1990. *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.

Pretty, J. 2003. "Social capital and the collective management of resources." *Science* 302:1912–1914.

Roth, A.E., V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir. 1991. "Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study." *The American Economic Review*, pp. 1068–1095.

Sethi, R., and E. Somanathan. 1996. "The evolution of social norms in common property resource use." *The American Economic Review*, pp. 766–788.

Sinha, S.K., and J. Talati. 2007. "Productivity impacts of the system of rice intensification (SRI): A case study in West Bengal, India." *Agricultural water management* 87:55–60.

Stoop, W.A., N. Uphoff, and A. Kassam. 2002. "A review of agricultural research issues raised by the system of rice intensification (SRI) from Madagascar: opportunities for improving farming systems for resource-poor farmers." *Agricultural Systems* 71:249–274.

Takahashi, K., and C.B. Barrett. 2014. "The System of Rice Intensification and its Impacts on Household Income and Child Schooling: Evidence from Rural Indonesia." *American Journal of Agricultural Economics*, pp. 269–289.

Voors, M.J., and E.H. Bulte. 2014. "Conflict and the evolution of institutions: Unbundling institutions at the local level in Burundi." *Journal of Peace Research* 51:455–469.

Zwane, A.P., J. Zinman, E. Van Dusen, W. Pariente, C. Null, E. Miguel, M. Kremer, D.S. Karlan, R. Hornbeck, X. Giné, et al. 2011. "Being surveyed can change later behavior and related parameter estimates." *Proceedings of the National Academy of Sciences* 108:1821–1826.