PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA

TESIS DE GRADO
MAGISTER EN ECONOMIA

Díaz Díaz, Ela Koritza

Diciembre, 2020

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA

TREATMENT EXTERNALITIES AND COMMUNITY DRIVEN DEVELOPMENT: EVIDENCE FROM NICARAGUA.

Ela Koritza Díaz Díaz

Comisión

Jeanne Lafortune y Eugenio Bobenrieth

Santiago, Diciembre de 2020

Treatment Externalities and Community-Driven Development: Evidence from Nicaragua.*

Ela Koritza Díaz Díaz†

December, 2020.

Abstract

Do social interactions with neighbors have an impact on participation in cooperative organizations? Using data from a representative sample of beneficiaries from a community-driven development program in Nicaragua, I exploit exposure to a random variation in the number of neighbors assigned to treatment from a field experiment to measure the impact of interactions. I also introduce a binary choice model with social interactions. The model predicts that ones' participation can be complementary or substitute from their neighbors. My results are congruent with strategic substitution. One more neighbor assigned to treatment within a 500-meter distance reduces participation in productive and community organizations by 2.34 and 0.8 percentage points, respectively. These results do not seem to be driven by anticipatory effects in the control group. Moreover, these impacts are larger for those who borrow their neighbors' inputs and have stronger ties with them.

*This manuscript was written at the Microeconomics Masters' Thesis Seminar, Department of Economics, PUC-Chile. I thank my advisors Jeanne Lafortune and Eugenio Bobenrieth for their continuous support, guidance, and insightful comments throughout this process. I also thank Pablo Celhay not only for allowing me to be part of this project and use the data for this research but also for his encouragement and support over the past years. I will also like to thank the World Bank and the MEFCA team for their excellent work in designing the experiment and collecting the data. Finally, I specially thank my family, Mario Rolando Diaz, Glorys Koritza Diaz, Mario Gabriel Diaz and Sonya Diaz for their support, and friends Diego Cussen, Gustavo Sepúlveda, and Andrea Herrera for their support, ideas, comments and proofreading. Any errors and omissions are my own.

†Comments to `ekdiaz@uc.cl`

1 Introduction

Community-based development (CBD) and its more intensive variant, community-driven development (CDD), have been one of the "work-horses" of international development organizations over the past decades. According to Mansuri and Rao (2004), CBD is an umbrella term for projects that actively include beneficiaries in their design and management, while CDD refers to community-based projects in which communities directly control key project decisions, such as investment funds management. The objective of incorporating beneficiaries in the decision-making process is that participation becomes self-initiated action and should lead to better-designed projects (Mansuri and Rao, 2012). Participation within communities would encourage individuals to behave more pro-socially, and the program could accomplish an expansion of pro-social norms without modifying them (Avdeenko and Gilligan, 2014). In the long term, gains in social capital might facilitate economic development or help sustain program impacts (Woolcock and Narayan, 2000).

Although this approach has many potential gains for development[1] and advocates that support it (Dongier et al., 2003; Narayan, 2002), it also has skeptics (Cooke and Kothari, 2001; Summers, 2001; Mosse, 2001). Reviews about the effect of programs with this vision in a variety of contexts show mixed results (Tanaka et al., 2006; Mansuri and Rao, 2012). So, there is still potential to continue research in the area and see if benefits from CBD or CDD programs out-weight challenges and risks of this strategy, such as practical, conceptual, and institutional issues (Tanaka et al., 2006). While there is increasing interest in research about elite capture (Platteau, 2004; Fritzen, 2007; Alatas et al., 2019), social capital building (Avdeenko and Gilligan, 2014; Nguyen and Rieger, 2017), or whether these programs can change the institutional environment (Casey et al., 2012), we know little about how connections between participants in these projects influence participation and its consequences on group formation or membership.

In this research, I study the impact of social interactions in the decision to participate in productive and community organizations in an agricultural CDD program. Given the social nature of groups formed by the intervention, participation may not only depend on how well an individual is connected but also on how influential or important these connections are. Social interactions are a crucial component of participation, since evidence regarding CDD programs finds that participants in civic activities tend to be wealthier, more educated, of higher social status (by caste and ethnic-

[1] According to Dongier et al. (2003), CDD is an effective mechanism for poverty alleviation. It can enhance sustainability, make poverty reduction more responsive to demand, build positive social capital, and make programs more inclusive of poor and vulnerable groups.

ity), male, and more politically connected than nonparticipants (Mansuri and Rao, 2012). In this sense, worse off individuals in terms of their networks might end up excluded or benefiting less from participation.

I use data from project PAIPSAN (Proyecto Apoyo para el Incremento de la Productividad, Seguridad Alimentaria y Nutricional en la Costa Caribe Nicaragüense), which aimed to reduce food insecurity among communities in this region. To achieve this objective, the program asked beneficiaries to form cooperative groups called *solidarity groups*. The purpose of these groups was twofold. First, to collect each recipient needs and construct an *Innovative Development Plan* (PDI in Spanish) that guided the programs' intervention. Second, to accelerate the implementation by purchasing and distributing inputs, organizing workshops, demonstrative activities, and consultation meetings, among others. The intervention consisted of delivering agricultural inputs (such as seeds, tools, or fertilizer) in kind, technical assistance, and training.

I measure the impact of social interactions as the spillover effect of neighbors in the treatment group on participation. To this end, I combine data on the geographic location of each farmer within the community and exogenous variation of participating in organizations coming from the random treatment assignment. I use a similar approach to other experimental studies (Miguel and Kremer, 2004; Dupas, 2014; Haushofer and Shapiro, 2016; De Mel et al., 2008). To support the empirical analysis, I introduce a binary choice model with social interactions. In this model, the influence of neighbors' participation can be positive or negative. If the relationship is positive, this implies that my participation is complementary to my neighbors'. If the opposite happens, it represents a substitution dynamic.

My results are consistent with a negative treatment externality, which indicate strategic substitution in participation. We can think about this strategic behavior as *free-riding* from neighbors' participation. For the first follow-up, exposure to one more neighbor assigned to treatment within a 500-meter distance decreased participation in productive organizations by 2.34 percentage points and in community organizations by 0.8 percentage points. For larger distances (1km and 1.5 km), these effects increase in magnitude. They also do for the second follow up. These results remain unchanged when I use pairs cluster bootstrap to estimate the standard errors. With the inclusion of community fixed effects, the results are robust for the participation in community organizations using neighbors within a 500-meter distance. However, the impacts on productive organizations decrease in magnitude and significance, while still remain negative. This means that differences across communities can explain part of the impact.

Results also show weak evidence that the relationship between one's participation and the number of neighbors assigned to treatment is non-linear. When their number is small, or they live very close to each other, the effect is either positive or zero. However, as the number of neighbors grows or the perimeter expands, the sign shifts to negative. Further analysis reveals that the negative sign is only present when local density is above the median. This suggests that the number of social interactions must be sufficiently large for strategic substitution to take place.

To assess if these effects come from strategic substitution or not, I tested for alternative explanations. One possibility is that they come from anticipatory effects in the control group. If the treated in the first round get sufficiently good outcomes due to the intervention, control group individuals might prefer to wait for their turn in the second round and reduce their participation in other organizations not promoted by the program. If this were to happen, we would see that as the number of treated individuals increases, participation in these organizations would decline. Results do not support this claim. The negative effects are concentrated in participation in solidarity groups for those in the treatment group.

Ruling out anticipatory effects, I explore several mechanisms that could explain why people would substitute participation with their neighbors using heterogeneous effects. First, I evaluate if people substitute by borrowing their neighbors' inputs or by learning about training sessions through them. As explained by Jackson (2010), this type of strategic substitution is called a *"best-shot"* public goods game. In this study's setting, since the program's management was mostly handled by the representatives from the organizations, if an individual wanted to collect their inputs or attend project-related activities, they had to be part of them (or have a friend who did). I find evidence to sustain this hypothesis. The coefficient of the number of neighbors assigned to treatment is larger for individuals that in need they would go to their neighbors for inputs or money. I also find negative externalities in attendance to training sessions, which means that people substitute assistance with their neighbors.

Second, I see whether sharing a common good (like the land) or lower levels of human capital (lower knowledge or undervaluation of benefits) makes the substitution more likely. These explanations are not backed by the data. For individuals with communal land property rights, the externality is not statistically different from zero as opposed to those with private land rights. Furthermore, I do not find that less educated or poorer individuals are more likely to interchange participation.

Although I find that less productive individuals substitute their participation with more productive neighbors, the most productive also do it, so it is not a behavior explained alone by low human capital.

Finally, I evaluate if the strategic substitution was heterogeneous in the nature of the bond with the neighbors. The results show that having stronger ties with neighbors or sharing characteristics such as gender or belonging to an indigenous group decreases the likelihood of participating in organizations. This means that stronger ties might be necessary to allow for strategic substitution to happen in this context.

Overall, the heterogeneity analysis reveals that the way individuals relate to their neighbors explains most of the observed patterns. The fact that the externality is larger when one can go to a friend in need or if we share a stronger tie is evidence of it. Also, the results for land-sharing are interesting on their own and indicate possible complementarities in production for those individuals (meaning that substitution could be costly for them).

This research contributes to the literature in several ways. First, about group membership in CBD interventions. To this day, studies have focused on determinants (Godquin and Quisumbing, 2005) and measuring the impact of programs that promoted group membership in different outcomes (Gugerty and Kremer, 2008; Deininger and Liu, 2009; Labonne and Chase, 2011). However, there are no studies about treatment externalities in these types of programs. These evaluations show that although people might benefit from micro-credit self-help groups (Deininger and Liu, 2009), outside funding might not expand organizational strength, and could weaken the role of the disadvantaged (Gugerty and Kremer, 2008). There are also possible crowding-out effects between civic activities (Labonne and Chase, 2011).

Second, to experimental evaluations with treatment externalities. Other studies find positive or zero treatment externalities of subsidies (Miguel and Kremer, 2004; Dupas, 2014; Haushofer and Shapiro, 2016; De Mel et al., 2008). In contrast, I observe a negative externality in a subsidized intervention. Third, to social strategic effects and free-riding literature. The free-rider problem has usually been studied in public goods provision (Bramoullé et al., 2007). However, I do not find that common-good sharing facilitates free-riding. Moreover, people in this program seems to free-ride from neighbors knowledge of training, which is similar to what Bandiera and Rasul (2006) find in their social learning framework.

2 Context

2.1 Program Description

In this research, I study participation in productive and community organizations by farmers from project PAIPSAN. This intervention was coordinated and implemented by the MEFCA (Ministerio de Economía Familiar, Comunitaria, Cooperativa y Asociativa) and its primary objective was to increase household food and nutritional security in rural communities in the region. The Caribbean Coast of Nicaragua is divided into two autonomous regions, the Northern (RACCN - Región Autónoma de la Costa del Caribe Nicaragüense) and the Southern (RACCS - Región Autónoma de la Costa del Caribe Nicaragüense Sur). The project covered more than 246 communities in 15 municipalities, eight are from the Northern (Waslala, Bonanza, Sinua, Rosita, Mulukukú, Waspan, Prinzapolka, and Puerto Cabezas) and seven from the Southern (La Cruz de Río Grande, Paiwas, El Tortuguero, Desembocadura de Río Grande, Laguna de Perlas, Kukra Hill, and Bluefields). As we can see in Figure 1, this area comprehends approximately 43% of the national territory (World Bank, 2019).

The government targeted this area due to the higher proportion of malnutrition in children, poverty, and indigenous population compared to the national levels. According to INIDE and MINSA (2014), in 2012, 23.3% of children under five from the Northern had stunting (the most severe form of malnutrition), which is 6 percentage points higher than the national average. Also, in 2014, the percentage of the citizens living under the poverty line in the entire region was 9.4 percentage points higher than the national average of 29.6%. On the other hand, in the same year, 35.29% of their population identified themselves as belonging to an indigenous group, which is more than three times the national level[2] (INIDE, 2016).

To select the beneficiaries, the ministry initially organized community meetings with local government authorities (indigenous council, elders, or neighbors associations). In them, they promoted the program and explained the objectives, eligibility criteria, and commitments. During these meetings, they identified possible participants who could be members of the assembly or those identified by other community members or local governments. Selection criteria were based on poverty status and occupation in line with the project objectives.

[2]The Caribbean Coast has been traditionally inhabited by the *Misquito or Miskitu* ethnic group, which is the largest ethnic minority in Nicaragua, representing 52.1% of the indigenous population.

The most important aspect of this program was the interventions' flexibility and the recipients' involvement. Beneficiaries were actively included in the program via public assemblies, PDI's formulation, and the program's implementation. The PDI's were the strategy created to intervene inside communities. They contained information regarding the type of initiatives the government was financing, budget, number of beneficiaries, dates, and short/medium-term goals, among others. After the ministry verified the eligibility of those interested to participate, and it was clear the beneficiaries pool, they asked recipients to form organized groups within their communities from 30 to 50 individuals, which they named solidarity groups. If by the time of the intervention, if there were already other organizations inside the communities, the ministry worked with them as well.

The *first* objective of these groups was creating the PDI's. For this, a second meeting was held inside the communities to collect the beneficiaries' proposals regarding the type of help they required. The aim was that the intervention was as flexible as possible and that it adjusted accordingly to individuals' needs. After proposals were gathered, a final version of the PDI was presented in each community so they could be legitimized and modified if necessary.

There were four PDI types: family agriculture, agro-industry, small-scale businesses and non-agriculture ventures and small scale fishery. Their components are summarized in Table 1. Although they were all different, they shared three basic features:

1. Financial support. It was delivered in kind and included seeds, tools, machinery, fertilizers, animals, etc.

2. Technical assistance. The ministry assigned an official who was in charge of handling all activities and follow up on the beneficiaries' progress of each PDI. In villages where they provided seeds, they carried demonstrative activities on how to set up plots or how to build family orchards. They also organized consultation meetings.

3. Training. The ministry also provided different types of training through workshops to complement financial investments. Farmers were instructed about crop diversification, fertilizer use, good agriculture practices, irrigation techniques, grain storage practices, how to increase sales, livestock care practices, among others. For agro-businesses, they promoted good manufacturing and hygiene practices. Fisher men also received training in hygiene practices. Since the ministry also sought to impact the beneficiaries' diet, they set up workshops about healthier food habits and meal preparations. In some cases, they also received training about the

importance and how to manage associative organizations. Training sessions usually lasted one day, but this varied according to the topic.

The *second* objective of these groups was to ease the programs' implementation. Each solidarity group chose representatives who had direct contact with the ministry officials. They helped to organize the workshops, demonstrative, and consultation activities. Representatives were also involved in the acquisition inputs and their distribution across communities. In practice, members of these groups also participated in the process evaluation of the program and their promotion to other village members.

Besides these management functions, the groups were conceived to be the first instance given any doubt or problem regarding the program. For PDI's execution, a system of "claims, complaints and suggestions" was created. If any recipient had a complaint or suggestion, they brought it first to their local representatives, and if they could not solve it, they involved the ministry. The idea was to handle problems locally and allow communities to do it.

Overall, these groups can be seen as a sort of "cooperatives", since individuals within them share goals in common, are managed by rules in which all members agree, and have a conflict resolution system. They also appear to provide benefits to participants. The most obvious comes from the program itself, such as agricultural inputs and capacity-building. But individuals could also beneficiate from collaboration, mutual support, network strengthening, and growth. However, they also impose a cost on them. The most obvious is the opportunity cost of time, which depends on individuals' socioeconomic status, employment, and family obligations. They also require effort and commitment from beneficiaries.

In agriculture, cooperatives can be a way to capture external economies of scale[3] and increase the market power of small-scale farmers concerning up and down-stream trading partners. (Valentinov, 2007). However, this program did not provide a link to markets which might be necessary to secure the long term benefits of associating (Mansuri and Rao, 2012), or any form of credit between beneficiaries.

[3] Johnson and Ruttan (1994) identify two economies of scale in agriculture: internal which are only related to the production process, and external, which are observed when large farms experience advantages in terms of access to inputs, credits, services and others.

2.2 Sample and Empirical Design

For the impact evaluation, a random sample of beneficiaries was drawn from the population. The sample size needed to detect impact was estimated to be 1,900 individuals, out of a total population of 14,000 people. Randomization was conducted at two levels: community and individual. Out of a total of 256 villages, 64 were offered to participate in the impact evaluation, and later, in each of them, program officials held assemblies to assign people to treatment and control groups through a public lottery. Treatment assignment was phased-in, meaning that both groups were offered to participate in the program. The program began for the treatment group in June 2017, while it did so for the control group in June 2018 (World Bank, 2019).

As we can see in Figure 2, baseline data were collected during April and May of 2017. Although it was estimated that a sample of 1,900 was needed, due to problems in the implementation of the baseline survey[4], the total number of households interviewed was 1,810, where 905 belong to the treatment group and 905 to the control group. After, two follow-ups were conducted. The first, a year after baseline, and the second one, during December of 2018 and March of 2019. The attrition in them is small. Of the 1,810 families in the baseline, 1,804 were interviewed in the first one, and in the second, 1,802 (World Bank, 2019).

3 Literature Review

3.1 CBD Programs with Group Membership

Due to the design of the intervention, this article relates to other CBD programs with group membership. For example, Deininger and Liu (2009) study the impact of micro-credit self-help groups in India. Their results show that participation in the program led to significant economic gains in the form of better nutrition and higher consumption levels as well as asset accumulation. Gugerty and Kremer (2008) examine a program that supported women's community associations in Kenya. The project's goal was to strengthen community organizations and improve agricultural practices and output. The authors find that little evidence that objective measures of group activity improved or that groups increased their external activities. However, outside funding did encourage the entry of younger, more educated women, employed in the formal sector, and people from outside the village. New entrants, men, and educated women assumed key leadership positions. Labonne and

[4]Two communities refuse to participate, letting out 80 families. Nine families were not found, and one refused to respond to the questionnaire.

Chase (2011) explore gains in the social capital of a project in the Philippines in which communities completed block grants for infrastructure investment. They find that the project increased participation in village assemblies and the frequency in which local officials meet with residents but had a negative impact on collective action.

These studies measure the impact of several CBD interventions with group membership in different outcomes. However, they do not focus on the effect that other participants within groups or networks may have in collective action or group participation. This research addresses this gap in the literature.

3.2 Treatment Externalities

This research is closely related to experimental studies that find positive externalities to other treated households who reside in its vicinity. For example, Angelucci and De Giorgi (2009) and Bobonis and Finan (2009) show that Progresa cash transfers can have behavioral effects among non-eligible households. Miguel and Kremer (2004) find that students that underwent intestinal worm infection, as well as children living within 3 kilometers of a school where most of the students were dewormed, experienced significant health-gains, despite not getting the treatment themselves. Dupas (2014) shows that the incidence of a subsidy for purchasing malaria bed nets has an independent effect on other households.

However, not all studies obtain large impacts or significant results. Haushofer and Shapiro (2016) find small or non-significant spillovers effects of unconditional cash transfers within villages and, although De Mel et al. (2008) find negative spillovers in profits from nearby enterprises of a firm subsidy, further analysis reveals that they were concentrated in the bamboo industry and not in the others.

Most of these studies have used geographical proximity to treated individuals to measure the spillover effects. Dupas (2014) uses the share of treated households within an arbitrary radius to other households, and Miguel and Kremer (2004) use the number of children within 3 kilometers from treated schools. Similar approaches were used in De Mel et al. (2008) and Haushofer and Shapiro (2016). In this research, since the share of treated households or the number of children within certain radio was exogenously determined by the random assignment, they can exploit this variation without running into the reflection problem (Manski, 1993). For my empirical analysis, I follow this same strategy.

Unlike the previous empirical work, I do not find evidence of positive or zero spillover effects, but rather, a negative treatment externality in participation. In my results, individuals reduce their involvement in organizations when the number of treated neighbors who live nearby gets sufficiently large enough, which indicate strategic behavior.

3.3 Free Rider Problem / Social Strategic Effects

This research is also related to free-rider and social strategic effects literature. The free-rider problem has usually been studied when there is a provision of a public good. For example, Bramoullé et al. (2007) investigate the incentives to provide non-excludable goods along with social or geographic links and find that it can lead to specialization, where some individuals contribute and other free-ride. However, this result could benefit society as a whole.

In my data, the incentive to free-ride does not seems to be driven by sharing a common good, but rather is only apparent between those who have excludable land rights. In this case, having communal land rights and sharing the land could make free-rider behavior to be more costly, reducing this incentive for this group.

Another area of research that has modeled this problem is social learning. Bandiera and Rasul (2006) demonstrate that farmers strategically delay the time they adopt a recommended farming technology to free ride on the knowledge generated by those in their neighborhood who experiment with it before them. In their results, they find that the relationship between technology adoption and the number of friends and family who adopt it is an *inverse-U*. The marginal effect of having one more adopter among friends and family is positive when there are few adopters in the network and negative when there are many.

Likewise, I find evidence of free-riding but not of an inverse U-shape relationship. In their case, people free-ride on their neighbors' knowledge, and although that in my context that also might be true, individuals also free-ride by exchanging agriculture inputs.

4 Theoretical Framework

To frame the empirical analysis, I now present a binary choice model with social interactions. In it, individuals must decide whether or not to participate in organizations. I assume that payoffs depend on individual participation decisions and the decisions made by their neighbors within the network.

There are information asymmetries: individuals do not have perfect information about their neighbors' type, which a preference for participating in organizations. In equilibrium, individuals decide based on the probabilities assigned to their neighbor types, conditional on their own type. This assumption is necessary for the equilibrium to be unique. Uniqueness allows me to apply Jackson (2010)'s definitions of strategic substitution or complementarity to my model, which is crucial to interpret the empirical results. In a world with complete information, it is possible to have multiple equilibria. In Galeotti et al. (2010), the authors explain that having complete information might result in multiple equilibria, a problem that can be solved with incomplete information. This also allows them to characterize network structure and define games as strategic complements and substitutes.

This model focuses on decision-making in the first stage of the evaluation, since when the control group receives the treatment, the analysis could get more complex[5]. However, in the empirical analysis, I also estimate the externalities for the second follow-up, but without separating between the two groups. Moreover, in the model, once individuals from the treatment group have made their choice about participating or no, they cannot repent from it[6].

4.1 Game Setting

Neighbors are players indexed by $i \in N \equiv \{1, ..., n\}$ and for simplicity, I assume that the position of each player was determined exogenously within the network [7]. Each community is a separated network, and two individuals i and j are connected if they live within an arbitrary distance d from each other. Let $Treat_i$ be a binary indicator of the random treatment assignment, $N^T_{d,-i}$ and $N_{d,-i}$

[5] People could care about what their treated neighbors did in the past, and about what their control neighbors are doing now.

[6] There are reasons to think this decision is irreversible. Individuals from the treatment group could not have received the intervention in the second stage of the program since it is against protocols. It is unlikely that individuals evade them since the program was heavily monitored. Even if they could, there may be costs of delaying entering, such as reputation for not cooperating. By themselves, this costs can be enough to prevent any delays in entry.

[7] This assumption could be relaxed assuming strategic network formation, which is more realistic, but makes the model more complex.

the number of neighbors assigned to the treatment group and the number of neighbors connected to i. Given her network, each player chooses simultaneously an action $y_i \in A \equiv \{0,1\}$, where 1 denotes taking the action, and 0 not taking the action. In this case, the action is participating in productive or community organizations. Let y_{-i} be the profile of actions of all players except i, X_i a vector of individual characteristics of i such as gender, age, ethnic group, educational level, etc., and X_{-i} a vector of characteristics of $i's$ neighbors.

Just before deciding, player i observes a vector of payments specific to an action, $\theta_i \equiv (\theta_{i0}, ..., \theta_{iK})$ which is private information (it is not observed by $j \neq i$). θ_i is an idiosyncratic preference for participation, which is different for each i. Also, let state variables X_i, and $N^T_{d,-i}$ associated with neighbor i to be of *common knowledge*. The payoff function V_{ik} of taking action $k \in A$ for player i is:

$$V_{ik}[Treat_i, y_{-i}(N^T_{d,-i}), X_i, X_{-i}, N_{d,-i}, \theta_i] = U_k(X_i, Treat_i) + \sum_{-i} W_k[y_{-i}(N^T_{d,-i}), X_i, X_{-i}, N_{d,-i}] + \theta_{ik} \quad (1)$$

where $U_k(\cdot)$ is the private utility of choosing k which depends on X_i, one's own characteristics and treatment assignment $Treat_i$. $W_k(\cdot)$ is the social utility of the election and measures the strategic effects on i's payoff from her friends' decision. Note that both functions depend on individual characteristics X_i and that social utility changes with the participation choices of its neighbors y_{-i}, neighbors characteristics X_{-i} and number of neighbors $N_{d,-i}$. Also, in this case $N^T_{d,-i}$ does not affect $W_k(.)$ directly, but rather, it serves as an instrument for y_{-i}.

It is important to highlight two aspects of the preference of participation θ_i are important to highlight. First, I suppose that they are private information. However, some neighbors may know the preferences of others (at least with some error). Even if this is true, this becomes untraceable in larger networks or in more urban areas, where density is higher. So it is reasonable to expect they are not observable for a proportion of the individuals. Second, although I assume that they are i.i.d in the theoretical model, I use clustered standard errors at a community level in the empirical analysis to allow these shocks to be potentially correlated.

4.2 Strategic Substitutes and Complements Definition

Before discussing the equilibrium of this game, I want to present two definitions that will be crucial for the interpretation of the empirical results: strategic complements and substitutes. According to Jackson (2010), a networks game exhibits *strategic complements* if it satisfies the property of *increasing differences*; that is for all c and $m \geq m'$:

$$u_c(1, m) - u_c(0, m) \geq u_c(1, m') - u_c(0, m') \tag{2}$$

where $u_c(.)$ is the payoff received by individual i of taking action 1 or 0, c is the number of connections an individual has, and m the number of players who take action 1. On the other hand, it exhibits *strategic substitutes* if it satisfies *decreasing differences* property; that is for all c and $m \geq m'$:

$$u_c(1, m) - u_c(0, m) \leq u_c(1, m') - u_c(0, m') \tag{3}$$

The way to interpret these definitions is that in a strategic complements game, the difference in the payoff of taking the action and not taking it is increasing in the number of neighbors who take the action, m. We can think about this difference as the incentive to take action 1. This means that as more neighbors choose 1, the incentive of choosing it becomes larger. The opposite happens with strategic substitution.

In terms of my model, these definitions are equivalent of taking the derivative of the difference between $V_{i1}(.)$ and $V_{i0}(.)$ against $N^T_{d,-i}$:

$$\frac{\partial[V_{i1}(.) - V_{i0}(.)]}{\partial N^T_{d,-i}} = \sum_{-i} \left[\frac{\partial[W_1(.) - W_0(.)]}{\partial y_{-i}}\right] \left[\frac{\partial y_{-i}}{\partial N^T_{d,-i}}\right] \lesseqgtr 0 \tag{4}$$

Expression (4) has two terms: the first measures the change in player i incentives to participate when her neighbors alter their participation decisions, and the second, the variation in the participation of i neighbors when the number of her neighbors assigned to the treatment group changes. Assuming that we hold control group participation constant, we would expect that the second term is always positive since treatment group participation should increase. However, the sign of the first term is ambiguous. Since y_{-i} is increasing, if social utility from participating is higher than of not participating, then it would be positive. The contrary would happen if the social utility from not participating is higher as opposed to doing it. This separates strategic complementarity from strategic substitution.

4.3 Bayes-Nash Equilibrium

This setting coincides with a *Bayesian Game*. In these games, at least one individual is uncertain about the payoffs or preferences of others, but other information about the game is of *common knowledge*. In my case, the only source of incomplete information is θ_i. The consequence of this failure is that we cannot solve the game using the typical Nash equilibrium. Harsanyi (1967) proposed a solution for the lack of information problem, in which although there is uncertainty about the players' type, each player can form prior believes about their distribution (e.g., knows the probability distribution of the types).

More formally, I assume that the players types $\{\theta_i\}_{i=1}^{I}$ are obtained from the prior distribution $p(\theta_1, ..., \theta_I)$, where θ_i belongs to the space Θ, which has a finite number of elements. Also, given the prior distribution, each player can use Bayes Rule to compute the conditional distribution $p(\theta_{-i}|\theta_i)$ of their neighbors types $\theta_{-i} = (\theta_1, ..., \theta_{i-1})$ from his type θ_i once they have received signals from their environment about their neighbors types. The set of pure strategies is Y_i, and the only two choices are to participate or not. A *pure strategy* for player i is the one that maps: $\Theta_i \to Y_i$, meaning her private information with the discrete choice Y_i. Also, given that player i knows his type, he evaluates his expected utility according to the conditional distribution $p(\theta_{-i}|\theta_i)$.

To characterize the equilibrium, I follow Xu (2018). Since i only cares about the differences in her expected payoffs of the actions, I normalize the average utility of action 0 making $U_0(x) = W_0(\ell, x, q) = 0$ for all $x \in S_x$, with $S_x \equiv (X_i', N_{d,-i}^T)$ be the set that contains all state variables, $\ell \in A$ and $q \in N$. Following the concept of Bayesian-Nash Equilibrium, the strategy of player i denoted by $r_i^*(\theta_i|S, \gamma)$, with $\gamma_k = (U_k, W_k)$ and $\gamma = (\gamma_1', ..., \gamma_k')$, maximizes her conditional expected payment given the other players equilibrium strategies $r_{-i}^*(\theta_{-i}|S, \gamma)$, that is:

$$r_i^*(\theta_i|S;\gamma) = \quad argmax_{k\in A} E\{V_{ik}[Treat_i, y_{-i}(N_{d,-i}^T), X_i, X_{-i}, N_{d,-i}, \theta_i]|S, \theta_i\}$$

$$r_i^*(\theta_i|S;\gamma) = \quad argmax_{k\in A}\left[U_k(X_i, Treat_i) + \sum_{\ell=0}^{K}\left\{W_k[\ell, y_{-i}(N_{d,-i}^T), X_i, X_{-i}, N_{d,-i}]\sum_{-i} p(r_{-i}^*(\theta_{-i}|S;\gamma)|S,\theta_i)\right\} + \theta_{ik}\right] \tag{5}$$

Moreover, Xu (2018) assumes that θ_{ik} is i.i.d. across actions and players and conform to an extreme value distribution with a density function $f(t) = e^t/1+e^t$, which is standard in this literature (Brock and Durlauf, 2002; Bajari et al., 2010). This assumption allows me to rewrite (5) in terms of equilibrium choice probabilities. Let $\sigma_{ik}^*(S;\gamma) = p(r_i^*(\theta_i|S,\gamma))$ and $\sigma_i^*(S;\gamma) = (\sigma_{i0}^*(S;\gamma), ..., \sigma_{ik}^*(S;\gamma))$

denote the equilibrium probabilities of action k and the profile of equilibrium actions, and also let $\sum(S;\gamma)^* = (\sigma_1(S,\gamma), ..., \sigma_n(S;\gamma))$ be the sequence of equilibrium profiles for all players. From (5) we have:

$$\sigma^*_{ik}(S;\gamma) = \frac{exp\left\{U_k(X_i, Treat_i) + \sum_{\ell=0}^{K}\left[W_k(\ell, y_{-i}(N^T_{d,-i}), X_i, X_{-i}, N_{d,-i}) * \sum_{-i}\sigma^*_{-i\ell}(S;\gamma)\right]\right\}}{1+\sum_{q=1}^{K} exp\left\{U_q(X_i, Treat_i) + \sum_{\ell=0}^{K}\left[W_q(\ell, y_{-i}(N^T_{d,-i}), X_i, X_{-i}, N_{d,-i}) * \sum_{-i}\sigma^*_{-i\ell}(S;\gamma)\right]\right\}} \tag{6}$$

Which is the typical form of a logit function except for equilibrium probabilities of the neighbors. Solving $\{r^*_1(\theta_i|S;\gamma)\ , ..., r^*_n(\theta_i|S;\gamma)\}$ is equivalent to solving $\{\sigma^*_1(S;\gamma), ..., \sigma^*_n(S;\gamma)\}$, according to Bajari et al. (2010). After characterizing the equilibrium, in addition to the assumption of extreme value distribution, Xu (2018) provides a second condition by which this equilibrium is unique (Lemma 1[8]). This condition involves making a restriction on the magnitude of the social effects, and it is given by:

$$\max_{k,m,\ell \in A} = |W_{k\ell} - W_{m\ell}| < \frac{K+1}{K} \tag{7}$$

Since that I am modeling a binary decision, this restriction means that the maximum influence of each farmer is bounded from above. This ensures weak independence, similar to the requirement that all roots lie outside of the unit circle in spatial autoregressive models. In my setting, this means that if the average effect of neighbors' participation increases in one percentage point, then $i's$ probability of participating has to change in less than one percentage point. Unfortunately, it is not possible to test for the validity of this assumption in these data. However, in previous applications to youth behavior, it usually holds (Sacerdote, 2001; Gaviria and Raphael, 2001; Kawaguchi, 2004; Carrell et al., 2008; Calvó-Armengol et al., 2009). Furthermore, in a similar setting to mine, González (2020) provides indirect evidence of the validity of this assumption using the parameters from a structural estimation to show that the elasticity of individual decisions with respect to network decisions is lower than one. But, this approach to test the veracity of this assumption goes beyond the scope of this research.

[8]The proof follows the contraction mapping argument.

5 Empirical Strategy

The model above suggests that there may exist a relationship between one's participation and the random assignment of one's neighbors. Empirically, I capture this by measuring the externality of neighbors from the treatment group in participation. For this, I exploit random variation in the local density exposure to individuals assigned to the treatment group created naturally by the experiment, using geo-referenced data from each household location inside the community. Due to imperfect treatment compliance, I rely on an Intention to Treat (ITT) to identify impact. The variation in local density can be considered exogenous since exposition to a certain number of neighbors assigned to treatment happens by "accident" as a consequence of random assignment in the baseline (since I control for the number of eligible neighbors). This approach has been widely used in this literature (Miguel and Kremer, 2004; Dupas, 2014; Haushofer and Shapiro, 2016; De Mel et al., 2008). The regression is of the form:

$$y_i = \beta_0 + \beta_1 Treat_i + \gamma_d N^T_{d,-i} + \phi_d N_{d,-i} + \alpha X_i + \epsilon_i \tag{8}$$

Here, y_i is a binary variable that takes the value of one if the individual participates in productive or community organizations and zero otherwise, $Treat_i$ is an indicator variable that takes the value of one if the individual was assigned to the treatment group in baseline and zero otherwise. In addition, $N^T_{d,-i}$ measures the number of neighbors assigned to treatment who live within distance d from i and $N_{d,-i}$ is number of eligible neighbors[9] who live within d. X_i is a vector individual characteristics of i.

My coefficient of interest is γ_d, and it measures the impact on $i's$ participation of having one more neighbor assigned to treatment within distance d, controlling for all neighbors that live within d. The average effect in treated individuals on the outcome is given by $\beta_1 + \gamma_d$, where the first term represents the *direct* effect of treatment assignment and the second the *externality*.

Although in equation (8) $N^T_{d,-i}$ enters linearly, there are reasons to think that this relationship is not linear. For example, early models of social influence on a binary decisions such as Granovetter (1978), state that others actions only matter if a certain number of people (or proportion) are also engaging the action. This type of behavior produces a non-linear relationship between the number of adopters and the probability of taking-the-action.[10] In my case, this could mean that

[9]This number includes the total number of neighbors within d, regardless if they are from the treatment or the control group, took the treatment, or refused to participate.

[10]In the original model, for each agent i there is a minimum proportion of $t_i \geq 0$, such that i adopts as soon as t_i

the relationship between y_i and $N^T_{d,-i}$ will be zero (or close) if $N^T_{d,-i}$ is below "i's threshold" and positive when above it. On the other hand, Bandiera and Rasul (2006) find an *U-shape* relationship between adoption of new farming technology and the number of friends or neighbors who adopted as well. I test for non-linearities using the following regression:

$$y_i = \beta_0 + \beta_1 Treat_i + f\left(N^T_{d,-i}\right) + \phi_d N_{d,-i} + \alpha X_i + \epsilon_i \tag{9}$$

For $f(.)$ I experiment with several specifications, such as splines and quadratic. I estimate equations (8) and (9) using a probit model and use clustered standard errors at the community level. I choose this model since, in approximately 50% of the observations of the participation in productive organizations, I have predicted probabilities below 20%, which is one of the thresholds where the linear approximation becomes less accurate and can produce probabilities out of range (Long, 1997). The choice between a probit and a logit model should not affect the results since these models provide very similar outcomes for the binary case. This election only seems to be relevant in the multivariate case (Hahn and Soyer, 2005).

I use clustered standard errors following Abadie et al. (2017) who claim that the use of clusters is a sampling or experimental design problem, rather than of correlation in residuals. In their paper, the authors prove that correlations between errors are neither a sufficient nor a necessary condition for the adjustment in standard errors to matter. Their use is justified if the treatment assignment or sampling process were made at a cluster level (leaving some clusters outside of the sample) and if there is enough variation in groups in the sample. In my data, for randomization, they first choose communities and later individuals. Therefore, some villages of the population are not present, and this justifies the use of clusters at the village level.

However, the use of clusters implies having several considerations regarding inference. First, not only the number of observations but also the number of clusters must go to infinity to have the desired asymptotic properties of the estimators (Wooldridge, 2010; Cameron and Miller, 2015). Second, there is a third dimension to consider: the number of observations by the cluster. The assumption that the number of clusters G must go to infinity presented in earlier work in the topic such as (Wooldridge, 2003, 2010; Cameron and Miller, 2015) assumes that the number of observations per cluster is fixed, and that might not be the case for many empirical applications. Moreover, unequal cluster sizes can have an impact on inference. As MacKinnon and Webb (2017) explains, inference

or more of the group has adopted.

using the cluster robust variance estimator can be unreliable even with 100 unbalanced clusters. Also, papers such as Bertrand et al. (2004), and Cameron et al. (2008) that can be used as a reference point regarding the appropriate G, use data sets primarily with equal-size clusters.

Hansen and Lee (2019) provide a useful framework to guide inference in clustered samples. In their paper, two assumptions are fundamental to have consistency, asymptotic normality, and efficiency. The maximum of the ratios $\frac{n_g}{n}$ and $\frac{n_g^2}{n}$ must converge to zero as $n \longrightarrow \infty$, with n being the total sample and n_g the number of observations per cluster. This implies that n_g is asymptotically negligible and that $G \longrightarrow \infty$. These assumptions are necessary for these asymptotic properties to hold.[11] It is possible that assumption 2 does not hold in my working sample,[12] which is necessary to have asymptotic normality and efficiency. Although it is not possible to test this empirically, I address this concern by using cluster bootstrapped standard errors.

One final remark regarding my empirical strategy is that the exogeneity in the number of neighbors assigned to treatment allows me to identify γ_d without worrying about some of the empirical problems social interactions literature has normally faced, such as reflection bias, unobservable correlated effects, and selection bias. The reflection bias, first mentioned by Manski (1993) refers to the fact that, if i influences j, then also j influences i, creating a multiplier effect. It is a simultaneous equation problem. The standard solution to it is to use an instrument. Unobserved correlated effects could be a problem if, for example, students within the same school have resembling performance because they come from a similar socioeconomic background or because they share the same teachers. This problem has led researchers to include fixed effects to absorb them. Finally, selection bias arises when peers share unobservable characteristics or proclivities that affect the outcome of interest. Efforts to eliminate this bias have focused on random peer assignments.

[11] See Theorems 10 and 11.

[12] I have that $n_g = 134$ and $n = 1672$, so ratio $\frac{n_g^2}{n}$ is $\frac{134^2}{1672} \approx 10.7392$

6 Results

6.1 Summary Statistics

Demographics. - For the data analysis, I restricted the sample to non-missing values of the variables in the three surveys, leaving a total of 1,672 observations out of an initial 1,810. In Table 2, I present, mean and standard deviation of demographic characteristics of the individuals and outcome variables in the baseline. We can see that a high percentage of recipients are illiterate (almost 30%), although the median age is about 39 years. Also, approximately 25% consider themselves as belonging to an indigenous group. The percentage of men and women in the sample is almost equal, slightly favoring women (they represent 51.26% of the observations).

Some of the dwelling characteristics reveal the poverty conditions of the beneficiaries: 50% of the individuals have dirt floors in their homes and, only 53.67% have electricity inside their homes. Also, almost all houses (90%) are made of wood walls. On the other hand, the primary occupations of program recipients are agriculture and cattle raising: 95.6% declare agriculture as a primary activity, and 79.55% engage in cattle raising. Only a small percentage is dedicated to forestry (0.6%). Additionally, we observe that production decision changes over seasons: 71.83% of the individuals engaged in sowing of any crop in *apante* (which is the most important growing season for this region) and 37.32% in *postrera* season.

Outcomes. - My empirical analysis focuses on the impact of treatment externalities in the participation in productive and community organizations, defined as discrete choice variables. Solidarity groups are a type of productive organization. For the majority of empirical analysis, I use participation in productive organizations in general. This is because the ministry worked with existing organizations (such as cooperatives, unions, or ONG's). Other types of groups, such as water and health committees, parents and church associations, and member of the indigenous government, will be in the community organizations category. Among the participants in my sample, very few people participate in any productive organization (less than 2%) in the baseline. However, around 17% participate in any community organization. The majority of this is driven by participation in religious organizations (which corresponds to 8.7%).

Number of neighbors. - Since I am interested in the impact of the number of neighbors assigned to treatment conditional on the number of eligible neighbors, this is equivalent to looking at the proportion of neighbors assigned to treatment within a certain radius. In Table (3), I present descriptive statistics of the share of neighbors assigned to treatment that live within radius d from i in baseline. I choose four different distances: 250-meter, 500-meter, 1km, 1.5 km, which I use in the empirical analysis.

We can observe significant dispersion in the share of treated, which is increasing as I expand the perimeter. In all distances, there are a fraction of individuals who were not exposed to another treated. However, this number decreases for the larger ones. The proportion of neighbors assigned to treatment living within a distance of 250-meter is on average 30.18% with a standard deviation of 25.66%. Its median is larger (35.29%), which means that the distribution is skewed to the left. The maximum proportion of treated for this distance is 87.5%. For 1.5km, the average share of treated is 43.53% with a standard deviation of 17.94%. As in the case of 250-meter, the median is above the mean (in 49.65%). For this extension, the maximum proportion of treated is 90%.

6.2 Empirical Strategy Validation

To validate the empirical strategy, in Table 4 I present linear regressions on pre-treatment outcomes and baseline covariates using equation (8) without additional controls. If the number of neighbors assigned to treatment is exogenous, we should not see that γ_d is significant for baseline variables. I look at farmers' demographics such as gender, age, literacy, and ethnicity; household characteristics like the number of household members, number of rooms in the house, and if the individual's dwelling has dirt floors and electricity. I also see some of the productive features of the beneficiaries, such as agriculture and cattle raising as a primary occupation, if she decides to produce, and corn and bean productivity in apante season.

I observe that the number of neighbors assigned to treatment conditional on the number of eligible is not correlated with any of the demographic characteristics of the farmers and baseline outcomes. However, it is correlated with household attributes such as the number of household members (positively) and the probability of having dirt floors (negatively) when using a 500-meter distance. Moreover, we can see that productive traits such as agriculture and cattle raising as a primary activity, and bean productivity in apante are also correlated with the number of treated within 500-meter and 1km.

I do not believe this is caused by the endogeneity of the number of neighbors assigned to treatment, since randomization produced two statistically comparable groups,[13] meaning that the treatment

[13]Table A.1 presents treatment balance in the baseline. Overall, randomization was performed correctly since I do

assignment is exogenous. As a result, the exposure to a determined number of neighbors assigned to treatment is also another source of exogenous variation. The correlations observed could be the result of bad luck (especially for dwelling characteristics since the variables are only significant for a 500-meter distance), and the existing patterns can be due to similarity in characteristics shared by farmers and possible spatial correlation of the variables.

Furthermore, to test whether these coefficients are were jointly significantly different from zero, I performed an F-test after running a seemingly unrelated regression (SUR). In this model, right-hand side variables are all the same but, one forms a system of equations of "seemingly unrelated" variables (that is, the origin of its name), regressed on an exogenous variable. The basic assumption of the model is that errors between equations are correlated for each i. For columns (2) to (4) of Table 4, I can reject the hypothesis that coefficients are jointly equal to zero. To account for this potential source of bias, I control for the variables in this table in my regressions.

6.3 Main results

In Table 5, I present the treatment and the number of neighbors assigned to treatment coefficients from equation (8) for 250-meter, 500-meter, 1km and 1.5km distances on each panel, respectively. In all regressions, I control for the number of eligible individuals within each radius. Columns (1) and (4) regress the participation in productive and community organizations without additional controls, in columns (2) and (5) I add baseline outcome as control, and in columns (3) and (6) individual characteristics from Table 4. The results from this table indicate that I cannot reject the hypothesis of no treatment externalities from neighbors within 250-meter, but I can reject they are different from zero for larger distances.

In column (1) of panel B, we observe that having one more neighbor assigned to the treatment group living within a 500-meter distance decreases participation in productive organizations by 2.3 percentage points. For community organizations, the externality is also negative, but its magnitude is smaller: we can see in column (4) that an increase in one neighbor assigned to treatment decreases participation by 0.8 percentage points. The size and significance of the coefficients are very similar when I control for baseline outcome and individual characteristics. Additionally, the coefficients on the number of treated neighbors increase for panels C and D (that show results for 1km and 1.5 km) and remain significant.

not observe systematic statistically significant differences between the two groups. Although there are differences in some variables, their significance is only at 10%.

It is difficult to assess whether one has a small number of clusters bias since there is not a "clear cut" about the number of clusters by which asymptotic properties hold. However, when faced with the problem of having a small number of clusters, the common solution is to use bootstrap. Clustered bootstrapping methods were first introduced by Cameron et al. (2008), and they usually work well unless one has a small number of treated or untreated clusters (MacKinnon and Webb, 2017). To perform the pairs cluster bootstrap as explained by Cameron et al. (2008), one must draw random samples of clusters from the original data and then form the standard errors or confidence intervals from the bootstrapped samples.

To see if results from Table 5 change when using bootstrap, in Table A.2, I show the results of the participation in organizations for the first follow up using pairs cluster bootstrap. We can see that the results are robust in this sense. Standard errors are very similar to the ones in the original sample, and only in a few cases, the significance decreases.

On the other hand, individuals may also worry about behaviors at more aggregated levels, such as their community. Also, there could be differences in how the program was implemented across villages. For this, I see if my results are robust to the inclusion of community fixed effects. I present them in Table A.3. We can see that the results hold for the participation in community organizations for a 500-meter distance, but not for productive organizations. In the former case, the coefficients decrease in magnitude and significance. However, they remain negative. This means that differences across communities can explain part of the impact.

In Table 3, I show the results from the second follow up. As in the case of the first follow up, I show the treatment and the number of neighbors assigned to treatment coefficients for the distances mentioned above and control for the number of eligible. The results indicate that the effect of the externality seems to get accented compared with the first follow-up. In column (1) of panel B, we see that having one more neighbor assigned to treatment within 500-meter decreases participation in productive organizations by 4.03 percentage points. This coefficient is 1.69 percentage points larger than the one found in Table 5 for that distance. For participation in community organizations, exposure to one more neighbor assigned to treatment decreases the probability to participate in 1.3 percentage points, which is 0.53 percentage points larger than for the first follow-up.

Although the impacts found in Tables 5 and 6 might seem small, they are not when we compared them to treatment take up. If I divide the two coefficients, the externality is about 13.28% of the treatment impact in participation in productive organizations for the 500-meter distance in the first follow up. This percentage increases to 21.28% for the 1.5 km radius. When doing this same exercise for the second follow-up, the magnitude of the externality against treatment is even larger.

The results so far are consistent with participation in community groups as being strategic substitutes. If we were interested in measuring the overall treatment effect, its magnitude would be smaller than take-up since this externality decreases participation.

6.4 Non-linearities

In what remains of the analysis, I focus on the effects on the first follow up. As discussed in Section 5, the relationship between participation and the number of neighbors assigned to treatment may not be linear. In Table 7, I show the results of equation (9) using splines. For each distance, I divided the number of neighbors assigned to treatment into quintiles. The omitted category in all panels is the first quintile, which is having zero neighbors for 250-meter and 500-meter, and having 0 to 1 neighbors for 1km. I find weak evidence of non-linearities.

For the participation in productive organizations, moving from having zero to one neighbor assigned to treatment (second quintile) for a 250-meter distance increases the probability of participating by 8.8 percentage points. The coefficients of the other quintiles for that distance are not significant. When using a 500-meter distance, only the fifth quintile is statistically different from zero, and the externality is now negative. For community organizations, I observe a similar pattern. The impact of neighbors living more closely (250-meter) is not statistically different from zero, and the negative impact comes from the quintiles with a larger number of treated neighbors (fourth and fifth).

To understand these results better, in Table A.4 I divide the average effect by the number of eligible neighbors above and below the median of the distribution. The evidence suggests that the externality is only negative when there is a higher density. When the number of eligible is below the median, the coefficient is not statistically significantly different from zero. However, the externality is negative and statistically significant for a 500-meter and 1km distance while the number of eligible is above the median.

Finally, to see if the shape of the relationship between participation and the number of treated is an inverted-U, in Table A.5 I show the results from a quadratic polynomial. I do not find that the relationship follows an inverted-U. The majority of the coefficients of the square number of neighbors are not significant or positive. And, the coefficient of the number of treated is always negative or not significant.

6.5 Alternative Explanations

6.5.1 Anticipatory Effects

The results from the previous section suggest that there is a strategic substitution with neighbors' participation. One potential alternative is that the negative effects are caused by anticipatory effects in the control group. If treated individuals in the first round get sufficiently good outcomes due to their participation in solidarity groups, their neighbors from the control group might prefer to wait for their "turn" to participate instead of seeking other sources of help in the short term. This type of conduct will not be congruent with strategic substitution.

The treatment group received the program in June 2017, which is before two crop seasons in Nicaragua, *postrera* and *apante*. Bean and maize can be cultivated in both of these seasons (Instituto Nacional Tecnológico, 2017). Since the control group would receive the program a year after the treatment, this gives them sufficient time to see the outcomes of their neighbors in the treatment group and make decisions regarding participation.

If this hypothesis is true, then I should see a *decrease* in the control groups' participation in other organizations (such as cooperatives, ONG's or unions) not promoted by the program in the first follow up. I present these results in Table 8. I exclude community organizations from the analysis. I do not find support for this claim in the results. None of the coefficients of the number of neighbors assigned to treatment are significant for the participation in other productive organizations, for both control and treatment groups. However, I see is that the negative effects come from substitution between treated individuals in the solidarity groups.

In column (4) of panel B, we can see that for the full sample, having one more neighbor assigned to the treatment group within a 500-meter radius decreases participation in solidarity groups by 1.85 percentage points. Moreover, from columns (5) and (6), it is evident that this is driven by treatment group behavior. In column (5), I observe that for those in the treatment group, exposure to one more treated neighbor decreases the probability of participating in solidarity groups by 3.23 percentage points. For the control group, the coefficient of the number of neighbors assigned to treatment is not significant.

6.6 Mechanisms

Ruling out that the negative impact comes from anticipatory effects in the control group, I now explore several mechanisms that could explain strategic substitution.

6.6.1 Best-shot public goods game?

Under a *best shot* public goods game, the incentive to engage in the action is lower if any of my neighbors do it before me. As explained by Jackson (2010), the typical case is information acquisition, but one could think about buying a book or a tool as possible actions. The key about this setting is that transaction costs are low, and that people have a higher utility if their neighbors do the action instead of them, since they do not pay for the cost.

How this type of game applies here? Since the programs' management was handled mostly by the community groups, if someone wanted to collect their inputs or attend training sessions, they had to be part of the organization (at least at some point in time). However, it is possible that if somebody has a neighbor from whom she can borrow inputs or learn about the training sessions or other project-related activities, she might prefer not to participate in them.

To test if strategic substitution can be explained by this, in Table 9, I present results from the participation in productive organizations using lending relationships in case of need and attendance to training sessions. I use two variables to measure lending relationships, if the person claims she would go to their friends or family for seeds or tools, and if she would go to them for money. I assume that if a person said they would turn to someone when they need to, they are more likely to substitute their participation by borrowing from neighbors. This is because they are the most predisposed to do it when faced with constraints (e.g., time). On the other hand, negative externalities in training sessions could also indicate strategic substitution in organizations.

I find that these conjectures are possible explanations of the results found. For those that claim they would go to their friends or family for inputs or money, the coefficient of the number of treated is larger. Also, having one more treated neighbor reduces the likelihood of attending training sessions.

In column (1) of panel A, I observe that for the ones that say they would go to their friends for seeds or tools, having one more treated neighbor to a 500-meter distance reduces their participation in productive organizations by 7 percentage points. However, the ones that report they would not go to their friends also free-ride but to a lesser extent. For them, one more treated neighbor decreases the probability to participate by 1.68 percentage points. The results for a 1km distance and money lending are very similar. In panel B, I estimated the treatment externalities for those who attended to at least one and more than two training sessions over the past month. We can see that for a 1km distance, one more neighbor assigned to the treatment group diminishes assistance at least one training session by 3.45 percentage points and attendance to more than two times by

2.06 percentage points.

These results suggest that a proportion of individuals free-ride from their neighbors' participation congruent a best-shot public goods game. In the subsequent sub-sections, I will address some other reasons why this behavior may take place.

6.6.2 Sharing a Common Good

Following the public goods argument, if two people share a common good, the incentive to free-ride might be larger because it is not excludable. In 2003, the Nicaraguan Government enacted law N° 445, which gave communal property rights to indigenous populations of the Atlantic Coast and Bocay, Coco, Indio, and Maiz rivers. The importance of this law is twofold. It created administrative figures within communities that are responsible for the administration of the land. And also, it forbids their selling to third parties. However, it permits their lending under the consultation of the communal authorities.

Although the indigenous government assigned to each family a certain number of hectares to use, they do not own the land. Not well-established property rights may facilitate coordination between neighbors and substitution in participation. Also, the exchange of information and agricultural inputs may be easier in this setting. So, I would expect that if the families share the land, their incentive to free-ride is larger.

To see to assess this, I allow the effect of treated neighbors to vary according to sharing the land with its neighbor. I defined land-sharing as i declaring to use communal land as well as her neighbor. Panel A of Table 10 reports results from equation (8) for individuals with communal and private land property rights. For those with communal property rights, I show two coefficients. One for the number of treated neighbors with whom i shares land and live to distance d, and for those with whom i do not share the land but live within distance d. In all regressions, I control for the number of eligible with whom i shares land or not, outcome, and individual characteristics at baseline.

I do not find that the negative social effects are driven by having communal property rights. For them, none of the coefficients of the number of neighbors assigned to treatment is significant. Instead, the externality is negative for beneficiaries with private land rights. In this case, having one more neighbor with private property rights within a 500-meter distance decreases their participation in productive organizations by 2.89 percentage points and in community organizations by 6.2 percentage points. This means that sharing land with neighbors does not explain free-riding.

6.6.3 Low Levels or Human Capital?

Since neighbors characteristics X_{-i} can make player i more or less likely to participate, individual and neighbors' education (or human capital) could be crucial in modifying agriculture decisions. For example, in technology adoption, Huffman (2001) argues that higher levels of education are associated with faster adoption in the short term and, Bandiera and Rasul (2006) find that more informed farmers are less sensitive to information about their network. Also, higher levels of education, wealth, and social status are associated with participation in civic activities (Mansuri and Rao, 2012).

Formal education improves the overall capacity to assimilate new information. More educated individuals could be more prone to participate since they could be more sensitive to the information about participation benefits when the program was promoted. Therefore, this could lead to less-educated having a lower (perceived) valuation of the organizations' benefits. If, as a consequence, they participate less, their incentive to free-ride from the most educated ones might be larger.

I use four criteria to sort farmers into high and low levels of human capital: literacy, poverty (having dirt floors), and agriculture productivity (maize and beans). If it is true that lower levels of human capital lead to an undervaluation of benefits, illiterate individuals who participate less as a result of this could engage in free-rider behavior when having a literate neighbor living nearby who does. The same may potentially happen in terms of poverty. Since agriculture is the only source of livelihood for most individuals in these villages, the least poor individuals may also be more talented and, hence, less inclined to free-ride, relative to others. An alike argument can be made for the most productive.

In general, my results suggest that low levels of human capital do not explain free-riding entirely. Although, in most cases, more disadvantaged individuals free-ride from their better-endowed neighbors, people with higher levels of human capital also free-ride from others who share this trait, and in some cases, from the ones in worst conditions.

In column (1) of panel B of Table 10, we can see that for the literate, having one more illiterate neighbor within a 500-meter distance decreases participation in productive organizations by 2.72 percentage points. Nevertheless, none of the coefficients for illiterate are significant for both productive and community organizations. Using poverty, I observe that for the not poor, exposure to one more neighbor with the same feature decreases participation in productive organizations by 1.91 percentage points. Coefficients for the poor are not significant in any of the columns.

In panels C and D, I present the results on productivity. For bean, I defined low productivity as reporting production below the median in apante. And for maize, as declaring having produced below the percentile 75. When using bean productivity, the results indicate that the more productive substitute their participation with others with and without the same attribute and less productive with their better-endowed neighbors. For the most productive, having another treated with the same trait living within 500-meter decreases participation in productive organizations by 3.5 percentage points and in community organizations by 0.7 percentage points. But also for them, one additional neighbor with lower productivity diminishes their involvement in community organizations by 1.6 percentage points. For the less productive, exposure to one more neighbor with the same characteristic decreases their likelihood to participate in community organizations by 2 percentage points.

For maize, in most cases, both the more and less productive substitute only from their neighbors without the same trait. In panel D, we can see that for those with higher productivity, exposure to one more neighbor with lower productivity decreases participation in productive organizations by 3.29 percentage points. On the other hand, for the less productive, an additional neighbor with the same attribute diminishes involvement in productive organizations by 2.13 percentage points and community organizations by 1.53 percentage points. Besides, one more neighbor without the same feature decreases their participation in community organizations by 5.16 percentage points.

To see if these results hold for the other distances, in Tables A.6 and A.7 show that when I extend the perimeter, free-riding from the less productive disappears and substitution from the more productive persist. This suggests that regardless of $i's$ productivity level if she has a more productive neighbor, her incentive to participate will be less.

6.6.4 Weak or Strong Ties

Throughout, I have taken neighbors to a certain distance to be the relevant reference group among which farmers' decisions to participate can be affected, as have done other in this literature (Miguel and Kremer, 2004; Dupas, 2014; De Mel et al., 2008; Haushofer and Shapiro, 2016). However, in reality, not all neighbors could be equally important when deciding participation, and this might be due to the nature of the tie. Although weaker ties might be responsible for spreading information across networks as Granovetter (1973) explains, strategic substitution may require stronger ties. This is because the person who free-rides *must* know she can go to her neighbor at any time to find out about the organization's activities and receive an "update".

I use two definitions for the strength of the ties, whether people know each other in baseline and sharing a characteristic such as gender and ethnicity. For each beneficiary, I have information about the people with whom she interacts the most within the community. With this, I can know if two people are acquaintances in the baseline. On the other hand, stronger ties may be greater in some groups due to homophily effects. Homophily is the tendency to associate and bond with those who are similar to me. This can affect behavior in a variety of ways. For example, it can affect workers' decisions of whether to drop out of the labor force (Jackson and Yariv, 2007), the speed of learning (Golub and Jackson, 2012), or generosity (Goeree et al., 2007).

I choose gender because recipient women may be driven to associate more with other women since the program had as a secondary objective to promote equal participation of women and increase their empowerment and leadership inside the communities (World Bank, 2019). Also, women involved in the promotion activities of the program may encourage other women's participation in their communities, and these women to others and suchlike. Generosity towards same-gender individuals might also be a factor affecting strategic substitution (Eckel and Grossman, 2001; Dufwenberg and Muren, 2006; Andreoni and Vesterlund, 2001; Ben-Ner et al., 2004). Regarding ethnicity, individuals may be obliged to relate more with those who share the same ethnic group due to similar customs, language or family ties.

My underlying assumption is that between people with stronger ties, strategic substitution is more likely to happen. I find the support of this claim in my results. It seems like the capacity to free-ride is larger when one has stronger ties with one's neighbor. This suggests that possessing strong ties with others may be a necessary condition for free-rider behavior to occur.

I present these results in Table 11. For those who know another beneficiary in the baseline, having one more treated neighbor whom they know within a 500-meter distance diminishes their participation in productive organizations by 14.16 percentage points. For them, the externality from those with weaker ties is also significant, but smaller: one more treated neighbor whom they do not know declines their participation in productive organizations by 4.31 percentage points. However, not knowing another beneficiary in the baseline does not impede free-riding. For those who do not know others in the baseline, an additional treated neighbor decreases their participation in productive organizations by 1.62 percentage points and community organizations by 0.67 percentage points.

In terms of shared characteristics, we can see that the externality is only negative for women. In column (2) of panel B, we observe that having one more female neighbor to a 500-meter distance decreases participation in productive organizations by 2.6 percentage points. Regarding ethnicity, I find that the substitution effects are only concentrated among indigenous populations. In column (1) of panel C, we observe that one more additional neighbor with the same ethnicity decreases participation in productive organizations by 3.4 percentage points and in community organizations by 0.8 percentage points.

7 Discussion and conclusion

This research aimed at estimating the impact of social interactions with neighbors in whether to participate in cooperative organizations. I find that people's involvement decreases with the number of neighbors (there is strategic substitution) and that these effects are concentrated amongst individuals exposed to a higher number of neighbors. I also find that stronger ties and being able to turn to a friend in case of need explain most of the results found.

Participation in civic activities is not an easy decision. Mansuri and Rao (2012) argue individuals embedded in their particular social groups and networks will balance all these costs and benefits before deciding to participate. However, the social costs can be prohibitively high for individuals in traditionally marginalized groups, such as women, members of disadvantaged castes, ethnic groups, or tribes. Moreover, they state that the decision to participate is not purely individual but also depends on groups' ability to come to a collective decision. In this sense, if individuals believe that the group will be ineffective or unable to reach consensus, they will be less inclined to participate.

In light of this, it is difficult to assert whether the results found in this research are undesirable. Although they indicate that people free-ride from their neighbors' participation, it is still unclear whether people found it not useful to participate or they had a higher opportunity cost. These two explanations lead to somewhat different policy implications. Unfortunately, I do not have the data to test any of them.

The results of this research also contribute to the discussion about the practical difficulties of conducting successful CDD programs. In his critique about how these interventions have been applied by the World Bank[14], Robert Chambers mentions several hidden externalities, like crowding out other new programs and diverting resources away from useful government services. On the other hand, these programs could suffer from elite capture, which may perpetuate society's existing

[14] Found in Kumar et al. (2005)

inequalities that the program was ostensibly trying to alleviate (Mansuri and Rao, 2012). Additional problems of these programs relate to the short-term horizon of the interventions (normally one year) and the long-term sustainability of the impacts (Tanaka et al., 2006).

Several policy lessons to prevent free-riding can be drawn from this research. Evidence suggests that only when projects link organizations with markets, or provide skills training, do they tend to improve cohesiveness and collective action beyond the life of the project (Mansuri and Rao, 2012). Although this program provided training, due to the heterogeneity of the topics addressed by communities, not all of them received the information regarding the importance and management of associative organizations. A more uniform delivery of information in this subject could have led to different results. Also, there was no "exit plan" at the end of the program to secure that organizations would endure over time, or link to agricultural markers, so it is possible that costs did not out-weight benefits for farmers. Furthermore, a longer time horizon might be necessary for people to understand the benefits of associating. For example, I find that farmers with community property rights do not free-ride, and their treatment coefficient is larger than for those with private property rights. This behavior may indicate that they already have internalized the benefits of cooperation before the intervention.

Finally, this study reveals the importance of social ties in shaping economic and social decisions, especially those that may not increase their well-being. People talk to one and another, ask advice from their friends and families, and engage in strategic behaviors. For this reason, it is crucial to think about externalities in general when designing a participatory interventions with group membership. In particular, they can undermine the net effect of the program.

References

Abadie, A., S. Athey, G. W. Imbens, and J. Wooldridge (2017). When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research.

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, R. Purnamasari, and M. Wai-Poi (2019). Does elite capture matter? local elites and targeted welfare programs in indonesia. In *AEA Papers and Proceedings*, Volume 109, pp. 334–39.

Andreoni, J. and L. Vesterlund (2001). Which is the fair sex? gender differences in altruism. *The Quarterly Journal of Economics 116*(1), 293–312.

Angelucci, M. and G. De Giorgi (2009). Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption? *American Economic Review 99*(1), 486–508.

Avdeenko, A. and M. J. Gilligan (2014). *International interventions to build social capital: evidence from a field experiment in Sudan*. The World Bank.

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010). Estimating static models of strategic interactions. *Journal of Business & Economic Statistics 28*(4), 469–482.

Bandiera, O. and I. Rasul (2006). Social networks and technology adoption in northern mozambique. *The Economic Journal 116*(514), 869–902.

Ben-Ner, A., F. Kong, and L. Putterman (2004). Share and share alike? gender-pairing, personality, and cognitive ability as determinants of giving. *Journal of Economic Psychology 25*(5), 581–589.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics 119*(1), 249–275.

Bobonis, G. J. and F. Finan (2009). Neighborhood peer effects in secondary school enrollment decisions. *The Review of Economics and Statistics 91*(4), 695–716.

Bramoullé, Y., R. Kranton, et al. (2007). Public goods in networks. *Journal of Economic Theory 135*(1), 478–494.

Brock, W. A. and S. N. Durlauf (2002). A multinomial-choice model of neighborhood effects. *American Economic Review 92*(2), 298–303.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009). Peer effects and social networks in education. *The Review of Economic Studies 76*(4), 1239–1267.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Cameron, A. C. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of human resources 50*(2), 317–372.

Carrell, S. E., F. V. Malmstrom, and J. E. West (2008). Peer effects in academic cheating. *Journal of human resources 43*(1), 173–207.

Casey, K., R. Glennerster, and E. Miguel (2012). Reshaping institutions: Evidence on aid impacts using a preanalysis plan. *The Quarterly Journal of Economics 127*(4), 1755–1812.

Cooke, B. and U. Kothari (2001). *Participation: The new tyranny?* Zed books.

De Mel, S., D. McKenzie, and C. Woodruff (2008). Returns to capital in microenterprises: evidence from a field experiment. *The quarterly journal of Economics 123*(4), 1329–1372.

Deininger, K. and Y. Liu (2009). *Longer-term economic impacts of self-help groups in India.* The World Bank.

Dongier, P., J. Van Domelen, E. Ostrom, A. Ryan, W. Wakeman, A. Bebbington, S. Alkire, T. Esmail, and M. Polski (2003). Community driven development. *World Bank Poverty Reduction Strategy Paper* (1).

Dufwenberg, M. and A. Muren (2006). Generosity, anonymity, gender. *Journal of Economic Behavior & Organization 61*(1), 42–49.

Dupas, P. (2014). Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica 82*(1), 197–228.

Eckel, C. C. and P. J. Grossman (2001). Chivalry and solidarity in ultimatum games. *Economic inquiry 39*(2), 171–188.

Fritzen, S. A. (2007). Can the design of community-driven development reduce the risk of elite capture? evidence from indonesia. *World Development 35*(8), 1359–1375.

Galeotti, A., S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv (2010). Network games. *The review of economic studies 77*(1), 218–244.

Gaviria, A. and S. Raphael (2001). School-based peer effects and juvenile behavior. *Review of Economics and Statistics 83*(2), 257–268.

Godquin, M. and A. R. Quisumbing (2005). Groups, networks, and social capital in rural philippine communities. *IFPRI. BASIS CRSP. October*.

Goeree, J. K., M. A. McConnell, T. Mitchell, T. Tromp, and L. Yariv (2007). Linking and giving among teenage girls. *Unpublished manuscript, California Institute of Technology*.

Golub, B. and M. O. Jackson (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics 127*(3), 1287–1338.

González, F. (2020). Collective action in networks: Evidence from the chilean student movement. *Journal of Public Economics 188*, 104220.

Granovetter, M. (1978). Threshold models of collective behavior. *American journal of sociology 83*(6), 1420–1443.

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology 78*(6), 1360–1380.

Gugerty, M. K. and M. Kremer (2008). Outside funding and the dynamics of participation in community associations. *American Journal of Political Science 52*(3), 585–602.

Hahn, E. D. and R. Soyer (2005). Probit and logit models: Differences in the multivariate realm. *The Journal of the Royal Statistical Society, Series B*, 1–12.

Hansen, B. E. and S. Lee (2019). Asymptotic theory for clustered samples. *Journal of econometrics 210*(2), 268–290.

Harsanyi, J. C. (1967). Games with incomplete information played by “bayesian” players, i–iii part i. the basic model. *Management science 14*(3), 159–182.

Haushofer, J. and J. Shapiro (2016). The short-term impact of unconditional cash transfers to the poor: experimental evidence from kenya. *The Quarterly Journal of Economics 131*(4), 1973–2042.

Huffman, W. E. (2001). Human capital: Education and agriculture. *Handbook of agricultural economics 1*, 333–381.

INIDE (2016). Encuesta de medición de nivel de vida. *Instituto Nacional de Información para el Desarrollo*.

INIDE and MINSA (2014). Encuesta de nicaragüense de demografía y salud 2011/2012. informe final. *Instituto Nacional de Información para el Desarrollo and Ministerio de Salud*.

Instituto Nacional Tecnológico (2017). Manual del protagonista: granos básicos.

Jackson, M. O. (2010). *Social and economic networks*. Princeton university press.

Jackson, M. O. and L. Yariv (2007). Diffusion of behavior and equilibrium properties in network games. *American Economic Review 97*(2), 92–98.

Johnson, N. L. and V. W. Ruttan (1994). Why are farms so small? *World development 22*(5), 691–706.

Kawaguchi, D. (2004). Peer effects on substance use among american teenagers. *Journal of Population Economics 17*(2), 351–367.

Kumar, N., A. Vajja, B. Pozzoni, and G. G. Woodall (2005). *The effectiveness of World Bank support for community-based and-driven development: An OED evaluation*. The World Bank.

Labonne, J. and R. S. Chase (2011). Do community-driven development projects enhance social capital? evidence from the philippines. *Journal of Development Economics 96*(2), 348–358.

Long, J. S. (1997). Regression models for categorical and limited dependent variables (vol. 7). *Advanced quantitative techniques in the social sciences*, 219.

MacKinnon, J. G. and M. D. Webb (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics 32*(2), 233–254.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies 60*(3), 531–542.

Mansuri, G. and V. Rao (2004). Community-based and-driven development: A critical review. *The World Bank Research Observer 19*(1), 1–39.

Mansuri, G. and V. Rao (2012). *Localizing development: Does participation work?* The World Bank.

Miguel, E. and M. Kremer (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica 72*(1), 159–217.

Mosse, D. (2001). People's knowledge, participation and patronage: operations and representations in rural development.

Narayan, D. (2002). *Empowerment and poverty reduction: A sourcebook*. The World Bank.

Nguyen, T. C. and M. Rieger (2017). Community-driven development and social capital: Evidence from morocco. *World Development 91*, 28–52.

Platteau, J.-P. (2004). Monitoring elite capture in community-driven development. *Development and change 35*(2), 223–246.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly journal of economics 116*(2), 681–704.

Summers, L. (2001). Speech at world bank country director's retreat. *Washington DC*.

Tanaka, S., J. Singh, and D. Songco (2006). A review of community-driven development and its application to the asian development bank. *Asian Development Bank*.

Valentinov, V. (2007). Why are cooperatives important in agriculture? an organizational economics perspective. *Journal of institutional Economics 3*(1), 55–69.

Woolcock, M. and D. Narayan (2000). Social capital: Implications for development theory, research, and policy. *The world bank research observer 15*(2), 225–249.

Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. *American Economic Review 93*(2), 133–138.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

World Bank (2019). Estudio de impacto de final del proyecto apoyo para el incremento de la productividad, seguridad alimentaria y nutricional de la costa del caribe nicaragüense. *Ministerio de Economía Familiar, Comunitaria, Cooperativa y Asociativa (MEFCA)*.

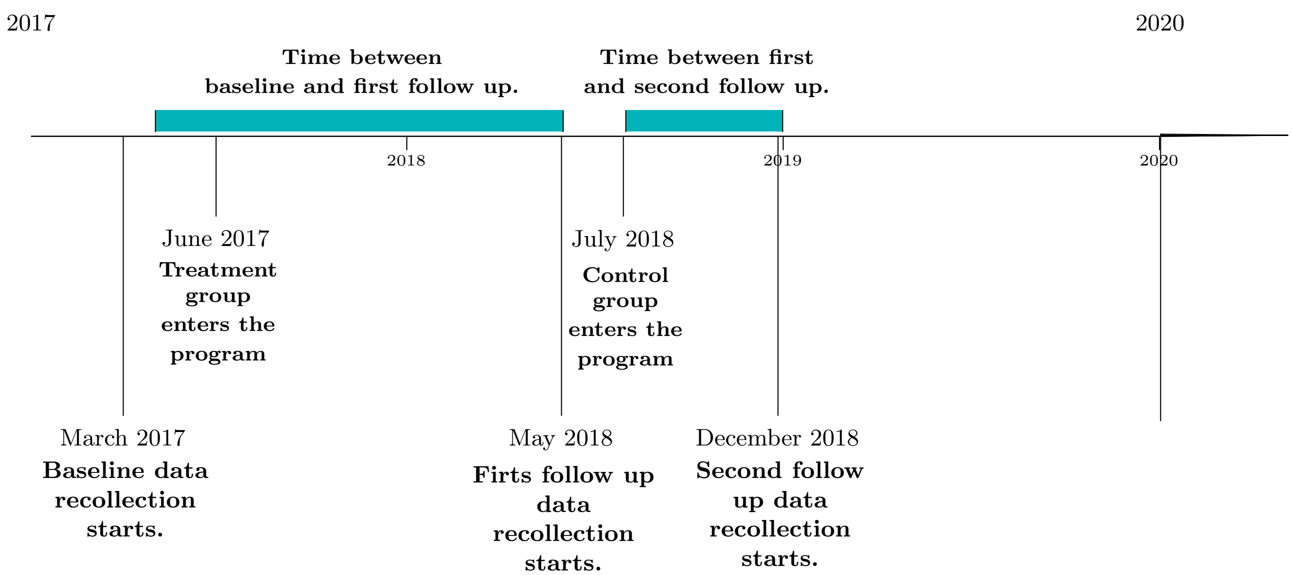
Xu, H. (2018). Social interactions in large networks: A game theoretic approach. *International Economic Review 59*(1), 257–284.

8 Figures

Figure 1: Geographic Incidence of the Project

Source: `https://www.economiafamiliar.gob.ni/paipsan-ccn/`

Figure 2: Project Timeline.



Note: Due to delays in data recollection in the first follow up, people from the control group received inputs up to one or two months later than what it was originally stipulated.

9 Tables

Table 1: Innovative Development Plans Types.

Sector	Objective	Help provided	Beneficiaries
Family agriculture	It had two sub-types: self-consumption, with the objective to increase production for self consumption and trade at a community level; and the other type oriented to sale of production surplus at a local level and seek profit.	Agriculture inputs such as seeds (maize, beans, cacao and coffee, roots and tubercles, musaceae, among others), fertilizer, tools and equipment, and in some cases small animals (chickens and pigs). Technical support for production and training.	11,000 in total, with 4,000 for self-consumption PDI and 7,000 for comercialization.
Agroindustry.	It supported different initiatives using cacao, maize and rice, among others.	Financial support for equipment, infrastructure and technical support for production and training.	500
Non-agriculture ventures and small rural businesses.	Strengthen market access and self employment opportunities.	Supplies, equipment and technical assistance and training for production, marketing and environmental management.	1,000
Small scale fishery.	It had two sub-types: self-consumption with the objective to increase household production for self consumption and to generate enough production to sell at a local level; and the second one for commercialization, aiming to specific markets.	Tools, equipment, machinery and technical support for production and training.	1,500 in total, with 1,000 for self-consumption and 500 to commercialization.

Source: adapted from (World Bank, 2019).

Table 2: Sample Characteristics in Baseline.

	Mean	Std. Dev.
Individual characteristics		
Age	38.6722	13.8954
Sex	0.4874	0.5000
Knows to read and write	0.7033	0.4569
Belongs to an indigenous group	0.2524	0.4345
Dwelling characteristics		
Household wall material: wood	0.8989	0.3015
Household floor material: dirt	0.4994	0.5001
Number of rooms in the house	2.6346	1.1818
Has electricity inside the house	0.5367	0.0809
Household owner	0.8846	0.3196
Productive characteristics		
Produces in apante season	0.7183	0.4500
Produces in postrera season	0.3732	0.4838
Corn productivity in apante season	10.1061	16.2246
Bean productivity in apante season	11.7595	18.3420
Primary activity: agriculture	0.9557	0.2057
Primary activity: cattle raising	0.7955	0.4035
Primary activity: forestry	0.0060	0.0771
Uses improved seed	0.1071	0.3093
Uses bio-fertilizers	0.1094	0.3123
Uses drip irrigation	0.0078	0.0879
Participation in organizations		
Productive	0.0173	0.1306
Community	0.1746	0.3798
Observations	1672	1672

Table 3: Variation in the Number of Neighbors Assigned to Treatment Conditional on the Number of Eligible in Baseline.

	Observations	Mean	Std. Dev.	Median	Min	Max
Share of treated within 250-meter	1668	0.3018	0.2566	0.3529	0	0.875
Share of treated within 500-meter	1668	0.3670	0.2399	0.4545	0	0.875
Share of treated within 1km	1668	0.4108	0.2043	0.4922	0	0.888
Share of treated within 1.5km	1668	0.4354	0.1794	0.4965	0	0.900

[1] Number of treated to each distance is the number of individuals assigned to treatment group in baseline different from i.

Table 4: Pre-Treatment Covariates Regressed on the Number of Neighbors Assigned to Treatment in Baseline at Different Distances.

	250-meter	500-meter	1 km	1.5 km
Individual characteristics				
Age	-0.1309	-0.1895	-0.1475	-0.1077
	(0.1385)	(0.1928)	(0.2110)	(0.2099)
Sex	-0.0004	0.0097	0.0052	0.0012
	(0.0053)	(0.0072)	(0.0094)	(0.0071)
Literacy	0.0020	0.0015	0.0051	-0.0059
	(0.0039)	(0.0044)	(0.0066)	(0.0044)
Indigenous	-0.0006	0.0081	-0.0005	-0.0145
	(0.0223)	(0.0147)	(0.0146)	(0.0127)
Participation in organizations				
Productive	-0.0007	0.0002	-0.0017	-0.0000
	(0.0013)	(0.0013)	(0.0016)	(0.0011)
Community	0.0071	0.0046	-0.0039	-0.0044
	(0.0086)	(0.0095)	(0.0089)	(0.0070)
Household characteristics				
Number of household members	-0.0046	0.0420**	-0.0109	-0.0175
	(0.0360)	(0.0205)	(0.0333)	(0.0299)
Number of rooms in the house	0.0034	0.0208	0.0409	0.0402
	(0.0267)	(0.0260)	(0.0383)	(0.0267)
Dirt floors	0.0049	-0.0196**	-0.0131	-0.0026
	(0.0130)	(0.0089)	(0.0139)	(0.0142)
Electricity	0.0093	0.0045	0.0142	0.0095
	(0.0098)	(.0149)	(0.0177)	(0.0157)
Productive characteristics				
Primary activity: agriculture	0.0037	0.0078*	0.0109***	0.0065**
	(0.0041)	(0.0039)	(0.0037)	(0.0030)
Primary activity: cattle raising	-0.0113	-0.0252**	-0.0273***	-0.0145
	(0.0109)	(0.0101)	(0.0081)	(0.0129)
Produces in apante season	0.0032	0.0089	-0.0055	-0.0134
	(0.0127)	(0.0140)	(0.0154)	(0.0104)
Corn productivity in apante	0.0355	0.1393	0.0833	-0.0389
	(0.2202)	(0.2991)	(0.2727)	(0.2609)
Bean productivity in apante	-0.6928*	-0.2828	-0.6274*	-0.4162
	(0.4043)	(0.4322)	(0.3730)	(0.3027)
Observations	1672	1672	1672	1672
Number of eligible control	Yes	Yes	Yes	Yes
P-value of joint significance test	0.3917	0.000	0.000	0.000

[1] Ordinary least squares coefficients of the number of neighbors assigned to treatment for each distance. Cluster standard errors at community level are shown in parenthesis.
[2] P-value of join significance test is the p-value of post estimation joint significance test of seemingly unrelated regression (SUR) model for variables in that column.
[3] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table 5: Treatment Externalities in Participation in the First Follow Up.

Dependent variables:	Participation in productive organizations			Participation in community organizations		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: radius of 250 meters						
Treatment	0.1777***	0.1777***	0.1758***	0.0278	0.0217	0.0188
	(0.0522)	(0.0521)	(0.0501)	(0.0193)	(0.0174)	(0.0175)
Number of treated to 250 m	-0.0035	-0.0035	-0.0018	-0.0025	-0.0041	-0.0029
	(0.0118)	(0.0118)	(0.0110)	(0.0057)	(0.0053)	(0.0048)
Panel B: radius of 500 meters						
Treatment	0.1761***	0.1761***	0.1750***	0.0267	0.0202	0.0177
	(0.0505)	(0.0504)	(0.0489)	(0.0189)	(0.0170)	(0.0173)
Number of treated to 500 m	-0.0234**	-0.0234**	-0.0216**	-0.0086**	-0.0098***	-0.0085***
	(0.0100)	(0.0100)	(0.0097)	(0.0039)	(0.0034)	(0.0028)
Panel C: radius of 1km						
Treatment	0.1676***	0.1676***	0.1692***	0.0247	0.0183	0.0167
	(0.0480)	(0.0480)	(0.0474)	(0.0187)	(0.0169)	(0.0173)
Number of treated to 1 km	-0.0328**	-0.0328**	-0.0309**	-0.0107**	-0.0103**	-0.0085**
	(0.0141)	(0.0141)	(0.0143)	(0.0046)	(0.0044)	(0.0043)
Panel D: radius of 1.5km						
Treatment	0.1602***	0.1602***	0.1617***	0.0251	0.0191	0.0176
	(0.0438)	(0.0437)	(0.0436)	(0.0183)	(0.0168)	(0.0173)
Number of treated to 1.5 km	-0.0341**	-0.0341**	-0.0343**	-0.0042	-0.0034	-0.0024
	(0.0143)	(0.0142)	(0.0133)	(0.0043)	(0.0042)	(0.0038)
Observations	1672	1672	1672	1672	1672	1672
Mean of dependent variable	0.1202	0.1202	0.1202	0.1166	0.1046	0.1046
Number of eligible control	Yes	Yes	Yes	Yes	Yes	Yes
Baseline outcome control	No	Yes	Yes	No	Yes	Yes
Individual controls	No	No	Yes	No	No	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[3] Dependent variables are indicator variables that take the value of 1 if the individual participates in productive or community organizations in the first follow up and zero otherwise. Independent variables are the treatment group indicator, that takes the value of one if the individual was assigned to treatment group in baseline and zero otherwise and the number of assigned to treatment neighbors to each survey participant at different arbitrary radius: 250 meters, 500 meters, 1km and 1.5km.
[3] All regressions are controlled by the total number of people to each respective radius, baseline outcome control is each outcome variable in baseline and individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[4] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table 6: Treatment Externalities in Participation in the Second Follow Up.

Dependent variables:	Participation in productive organizations			Participation in community organizations		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: radius of 250 meters.						
Treatment	0.1112**	0.1116**	0.1121**	0.0220*	0.0184	0.0187
	(0.0437)	(0.0436)	(0.0437)	(0.0121)	(0.0125)	(0.0123)
Number of treated to 250 m	-0.0188	-0.0188	-0.0168	-0.0103***	-0.0110***	-0.0114***
	(0.0125)	(0.0125)	(0.0106)	(0.0034)	(0.0030)	(0.0037)
Panel B: radius of 500 meters.						
Treatment	0.1079**	0.1082**	0.1102***	0.0200	0.0163	0.0169
	(0.0423)	(0.0422)	(0.0414)	(0.0126)	(0.0131)	(0.0130)
Number of treated to 500 m	-0.0403***	-0.0403***	-0.0375***	-0.0139***	-0.0144***	-0.0147***
	(0.0115)	(0.0115)	(0.0113)	(0.0031)	(0.0032)	(0.0032)
Panel C: radius of 1km.						
Treatment	0.0981**	0.0984**	0.1033**	0.0176	0.0138	0.0143
	(0.0445)	(0.0445)	(0.0430)	(0.0122)	(0.0125)	(0.0125)
Number of treated to 1 km	-0.0507***	-0.0506***	-0.0484***	-0.0135***	-0.0134**	-0.0140***
	(0.0175)	(0.0176)	(0.0168)	(0.0050)	(0.0052)	(0.0049)
Panel D: radius of 1.5 km.						
Treatment	0.0938**	0.0941**	0.0976**	0.0163	0.0127	0.0130
	(0.0447)	(0.0447)	(0.0432)	(0.0121)	(0.0124)	(0.0123)
Number of treated to 1.5 km	-0.0395*	-0.0395*	-0.0418**	-0.0110**	-0.0107**	-0.0109**
	(0.0213)	(0.0213)	(0.0187)	(0.0048)	(0.0050)	(0.0049)
Observations	1672	1672	1672	1672	1672	1672
Mean of dependent variable	0.2392	0.2476	0.2476	0.0601	0.0601	0.0601
Number of eligible control	Yes	Yes	Yes	Yes	Yes	Yes
Baseline outcome control	No	Yes	Yes	No	Yes	Yes
Individual controls	No	No	Yes	No	No	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[3] Dependent variables are indicator variables that take the value of 1 if the individual participates in productive or community organizations in the second follow up and zero otherwise. Independent variables are the treatment group indicator, that takes the value of one if the individual was assigned to treatment group in baseline and zero otherwise and the number of assigned to treatment neighbors to each survey participant at different arbitrary radius: 250 meters, 500 meters, 1km and 1.5km.
[3] All regressions are controlled by the total number of people to each respective radius, baseline outcome control is each outcome variable in baseline and individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[4] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table 7: Non-linearities Between Participation and the Number of Neighbors Assigned to Treatment.

Dependent variables:	Participation in productive organizations		Participation in community organizations	
	(1)	(2)	(3)	(4)
Panel A: 250 meters radius				
Treatment	0.173***	0.170***	0.028	0.023
	(0.052)	(0.051)	(0.020)	(0.020)
1 neighbour	0.082*	0.088**	-0.004	-0.001
	(0.043)	(0.042)	(0.023)	(0.022)
2 neighbours	0.035	0.037	0.017	0.013
	(0.042)	(0.042)	(0.026)	(0.027)
3-10 neighbours	0.037	0.051	-0.024	-0.023
	(0.049)	(0.048)	(0.028)	(0.026)
11-32 neighbours	-0.158	-0.133	-0.067	-0.065
	(0.114)	(0.104)	(0.088)	(0.079)
Panel B: radius of 500 meters				
Treatment	0.1726***	0.1704***	0.0285	0.0229
	(0.0502)	(0.0485)	(0.0185)	(0.0181)
1-2 neighbors	0.0331	0.0321	-0.0276	-0.0254
	(0.0346)	(0.0346)	(0.0236)	(0.0238)
3-5 neighbors	0.0263	0.0336	-0.0297	-0.0229
	(0.0492)	(0.0494)	(0.0270)	(0.0259)
6-26 neighbors	-0.0293	-0.0152	-0.0896***	-0.0835***
	(0.0729)	(0.0738)	(0.0245)	(0.0232)
27-57 neighbors	-0.3207*	-0.2855	-0.2423***	-0.2285***
	(0.1793)	(0.1843)	(0.0598)	(0.0648)
Panel C: radius of 1km				
Treatment	0.1741***	0.1748***	0.0267	0.0218
	(0.0508)	(0.0497)	(0.0194)	(0.0191)
2-4 neighbors	0.0466	0.0409	-0.0022	0.0028
	(0.0373)	(0.0359)	(0.0248)	(0.0239)
5-8 neighbors	0.0185	0.0118	-0.0002	0.0035
	(0.0542)	(0.0533)	(0.0340)	(0.0337)
9-50 neighbors	0.0421	0.0293	-0.0361	-0.0257
	(0.1098)	(0.1045)	(0.0587)	(0.0558)
51-72 neighbors	0.0444	-0.0206	-0.0285	-0.0089
	(0.3043)	(0.2814)	(0.1687)	(0.1524)
Observations	1672	1672	1672	1672
Mean of dependent variable	0.1202	0.1202	0.1046	0.1046
Number of eligible control	Yes	Yes	Yes	Yes
Baseline outcome control	No	Yes	No	Yes
Individual controls	No	Yes	No	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[3] Dependent variable is an indicator variable that takes the value of 1 if the individual participates in productive in the first follow up and zero otherwise. Independent variables are the treatment group indicator, that takes the value of one if the individual was assigned to treatment group in baseline and zero otherwise and the number of assigned to treatment neighbors to each survey participant at different arbitrary radius: 250 meters, 500 meters, 1km and 1.5km.
[3] All regressions are controlled by the total number of people to each respective radius, baseline outcome control is each outcome variable in baseline and individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[4] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table 8: Anticipatory Effects in the Control Group.

Dependent variables:	Participation in Other Productive Organizations			Participation in Solidarity Groups		
	Total (1)	Treated (2)	Control (3)	Total (4)	Treated (5)	Control (6)
Panel A: radius of 250 meters						
Treatment	0.0171			0.1348***		
	(0.0108)			(0.0373)		
Number of treated to 250 m	-0.0001	0.0013	-0.0023	-0.0028	-0.0154	0.0075
	(0.0026)	(0.0033)	(0.0052)	(0.0085)	(0.0102)	(0.0079)
Panel B: radius of 500 meters						
Treatment	0.0171			0.1334***		
	(0.0108)			(0.0354)		
Number of treated to 500 m	-0.0002	0.0019	-0.0016	-0.0185***	-0.0323***	-0.0027
	(0.0026)	(0.0036)	(0.0028)	(0.0069)	(0.0096)	(0.0051)
Panel C: radius of 1km						
Treatment	0.0175			0.1279***		
	(0.0110)			(0.0332)		
Number of treated to 1 km	0.0015	0.0028	0.0014	-0.0258**	-0.0447**	-0.0047
	(0.0022)	(0.0031)	(0.0019)	(0.0125)	(0.0199)	(0.0067)
Panel D: radius of 1.5 km						
Treatment	0.0176			0.1224***		
	(0.0112)			(0.0318)		
Number of treated to 1.5 km	0.0017	0.0020	0.0022	-0.0265**	-0.0412**	-0.0064
	(0.0016)	(0.0025)	(0.0014)	(0.0123)	(0.0182)	(0.0073)
Observations	1672	840	832	1672	840	832
Mean of dependent variable	.	0.0333	0.0180	.	0.2143	0.0805
Number of eligible control	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[3] Dependent variable is an indicator variable that takes the value of 1 if the individual participates in productive in the first follow up and zero otherwise. Independent variables are the treatment group indicator, that takes the value of one if the individual was assigned to treatment group in baseline and zero otherwise and the number of assigned to treatment neighbors to each survey participant at different arbitrary radius: 250 meters, 500 meters, 1km and 1.5km.
[3] All regressions are controlled by the total number of people to each respective radius, and individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[4] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table 9: Best-shot Public Goods Game?

Dependent variable: participation in productive organizations	Would go to his friends for seeds or tools in baseline		Would go to his friends for money in baseline	
	(1) Yes	(2) No	(3) Yes	(4) No
Panel A: Lending relationships in case of need				
Treatment	0.0702	0.1856***	0.0025	0.1911***
	(0.0640)	(0.0462)	(0.0581)	(0.0462)
Number of treated to 500 m	-0.0700***	-0.0168**	-0.0633***	-0.0171**
	(0.0119)	(0.0083)	(0.0125)	(0.0082)
Treatment	0.0889	0.1810***	0.0328	0.1864***
	(0.0691)	(0.0448)	(0.0581)	(0.0449)
Number of treated to 1 km	-0.0759***	-0.0233*	-0.0756***	-0.0227*
	(0.0209)	(0.0125)	(0.0201)	(0.0123)
Observations	138	1534	147	1525
Number of eligible control	Yes	Yes	Yes	Yes
Baseline outcome control	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

Dependent variables:	Attended at least one training session over the past month		Attended more than two training sessions over the past month	
	(1)	(2)	(3)	(4)
Panel B: externalities in training sessions				
Treatment	0.1608***	0.1497***	0.0965***	0.0890***
	(0.0432)	(0.0381)	(0.0360)	(0.0329)
Number of treated to 500 m	-0.0125	-0.0140	-0.0057	-0.0062
	(0.0115)	(0.0090)	(0.0060)	(0.0046)
Treatment	0.1496***	0.1438***	0.0868***	0.0836***
	(0.0366)	(0.0339)	(0.0272)	(0.0268)
Number of treated to 1 km	-0.0345***	-0.0361***	-0.0206***	-0.0216***
	(0.0131)	(0.0118)	(0.0065)	(0.0053)
Observations	1672	1672	1672	1672
Number of eligible control	Yes	Yes	Yes	Yes
Baseline outcome control	No	Yes	No	Yes
Individual controls	No	Yes	No	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[2] Individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[4] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table 10: Incentives to Free-Ride by Neighbors Characteristics (500-meter Distance).

Dependent variables:	Participation in productive organizations		Participation in community organizations	
	(1) Communal	(2) Private	(3) Communal	(4) Private
Panel A: communal property rights				
Number of treated with same same land	-0.0032		-0.0014	
	(0.0052)		(0.0027)	
Number of treated without without same land	-0.1058	-0.0289***	-0.0129	-0.0622***
	(0.1215)	(0.0093)	(0.0223)	(0.0144)
Observations	293	1379	293	1379
	Literate (1)	Illiterate (2)	Literate (3)	Illiterate (4)
Panel B: literacy				
Number literate treated	-0.0113	-0.0221	0.0008	-0.0155
	(0.0085)	(0.0151)	(0.0029)	(0.0097)
Number illiterate treated	-0.0272*	-0.0209	-0.0126	-0.0150
	(0.0161)	(0.0235)	(0.0082)	(0.0235)
Observations	1176	496	1176	496
	Not poor (1)	Poor (2)	Not poor (3)	Poor (4)
Panel C: poverty				
Number of not poor treated	-0.0191*	-0.0241	-0.0067	-0.0066
	(0.0109)	(0.0207)	(0.0063)	(0.0077)
Number of poor treated	-0.0325	-0.0097	0.0048	-0.0115
	(0.0354)	(0.0169)	(0.0302)	(0.0091)
Observations	837	835	837	835
	More productive (1)	Less productive (2)	More productive (3)	Less productive (4)
Panel D: bean productivity				
Number of more productive treated	-0.035**	-0.006	-0.007*	-0.019
	(0.017)	(0.014)	(0.004)	(0.005)
Number of less productive treated	-0.046	-0.014	-0.016*	-0.020**
	(0.029)	(0.018)	(0.009)	(0.010)
Observations	835	837	835	837
	More productive (1)	Less productive (2)	More productive (3)	Less productive (4)
Panel E: maize productivity				
Number of more productive treated	-0.0436	-0.0102	0.0188	-0.0516*
	(0.0268)	(0.0283)	(0.0272)	(0.0147)
Number of less productive treated	-0.0329**	-0.0213*	-0.0035	-0.0153***
	(0.0162)	(0.0120)	(0.0089)	(0.0047)
Observations	509	1163	509	1163
Treatment indicator	Yes	Yes	Yes	Yes
Number of eligible control	Yes	Yes	Yes	Yes
Baseline outcome control	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[2] Individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[2] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table 11: Strength of Ties (500-meter Distance).

Dependent variables:	Participation in productive organizations		Participation in community organizations	
	Knows another beneficiary in baseline (1)	Does not knows another beneficiary in baseline (2)	Knows another beneficiary in baseline (3)	Does not knows another beneficiary in baseline (4)
Panel A: village acquaintances				
Number of treated with strong ties	-0.1416***		-0.0147	
	(0.0463)		(0.0413)	
Number of treated with weak ties	-0.0431***	-0.0162**	-0.0117	-0.0067**
	(0.0149)	(0.0078)	(0.0104)	(0.0032)
Observations	238	1434	238	1434
	Males (1)	Females (2)	Males (3)	Females (4)
Panel A: gender				
Number of treated from the same gender	-0.0104	-0.0277***	0.0016	-0.0056
	(0.0218)	(0.0097)	(0.0086)	(0.0058)
Number of treated not from the same gender	-0.0180	-0.0138	-0.0056	-0.0021
	(0.0117)	(0.0132)	(0.0056)	(0.0081)
Observations	815	857	815	857
	Indigenous (1)	Non indigenous (2)	Indigenous (3)	Non indigenous (4)
Panel B: ethnicity				
Number of treated same ethnicity	-0.0337***	-0.0100	-0.0088***	-0.0046
	(0.0093)	(0.0144)	(0.0031)	(0.0068)
Number of treated not same ethnicity	0.0226**	-0.0441***	0.0062	-0.0179***
	(0.0115)	(0.0075)	(0.0051)	(0.0046)
Observations	422	1250	422	1250
Treatment indicator	Yes	Yes	Yes	Yes
Number of eligible control	Yes	Yes	Yes	Yes
Baseline outcome control	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis. Dependent variables are participation in productive and community organizations in the first follow up.
[3] Individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[4] *** Significant at 1%, ** significant at 5% and * significant at 10%.

A Supplementary Tables

Table A.1: Treatment Balance.

	250-meter	500-meter	1 km	1.5 km
Individual characteristics				
Age	0.4494	0.4303	0.4060	0.3832
	(0.7077)	(0.7036)	(0.7119)	(0.7216)
Sex	-0.0477*	-0.0474*	-0.0457	-0.0461*
	(0.0279)	(0.0277)	(0.0277)	(0.0274)
Literacy	0.0279	0.0277	0.0278	0.0243
	(0.0256)	(0.0259)	(0.0264)	(0.0266)
Indigenous	0.0344	0.0301	0.0279	0.0209
	(0.0251)	(0.0255)	(0.0213)	(0.0211)
Participation in organizations				
Productive	-0.0034	-0.0035	-0.0038	-0.0034
	(0.0063)	(0.0062)	(0.0062)	(0.0062)
Community	0.0385	0.0393	0.0388	0.0381
	(0.0232)	(0.0237)	(0.0234)	(0.0229)
Household characteristics				
Number of household members	0.0834	0.0775	0.0736	0.0666
	(0.1162)	(0.1148)	(0.1128)	(0.1113)
Number of rooms in the house	0.0035	0.0035	0.0105	0.0147
	(0.0642)	(0.0648)	(0.0647)	(0.0651)
Dirt floors	-0.0080	-0.0046	-0.0058	-0.0020
	(0.0211)	(0.0215)	(0.0204)	(0.0202)
Electricity	0.0444*	0.0459*	0.0479*	0.0473*
	(0.0264)	(0.0267)	(0.0260)	(0.0276)
Productive characteristics				
Primary activity: agriculture	-0.0168*	-0.0163	-0.0144	-0.0141
	(0.0100)	(0.0102)	(0.0097)	(0.0095)
Primary activity: cattle raising	0.0327	0.0312	0.0265	0.0273
	(0.0270)	(0.0274)	(0.0271)	(0.0265)
Produces in apante season	-0.0041	-0.0048	-0.0054	-0.0082
	(0.0272)	(0.0274)	(0.0271)	(0.0278)
Corn productivity in apante	-0.2065	-0.2098	-0.1724	-0.1758
	(0.6918)	(0.6975)	(0.6735)	(0.6624)
Bean productivity in apante	-0.0427	-0.2022	-0.3303	-0.3474
	(0.7886)	(0.7822)	(0.7908)	(0.7897)
Observations	1672	1672	1672	1672

[1] Ordinary least squares coefficients of treatment indicator for each distance. Cluster standard errors at community level are shown in parenthesis.
[4] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table A.2: Bootstrap Standard Errors for the First Follow Up.

Dependent variables:	Participation in productive organizations		Participation in community organizations	
	Below the median (1)	Above the median (2)	Below the median (3)	Above the median (4)
Panel A: radius of 250 meters				
Treatment	0.1049***	0.2469***	0.0082	0.0254
	(0.0258)	(0.0830)	(0.0233)	(0.0287)
Number of treated to 250 m	0.0239	-0.0040	-0.0001	-0.0058
	(0.0258)	(0.0138)	(0.0209)	(0.0049)
Observations	917	755	917	755
Mean of dependent variable	0.1099	0.1332	0.1078	0.1005
Panel B: 500 meters				
Treatment	0.1007***	0.2579***	0.0160	0.0159
	(0.0260)	(0.0823)	(0.0231)	(0.0264)
Number of treated to 500 m	0.0026	-0.0258***	-0.0098	-0.0077***
	(0.0187)	(0.0098)	(0.0153)	(0.0027)
Observations	924	748	924	748
Mean of dependent variable	0.1318	0.1045	0.1046	0.1045
Panel C: radius of 1km				
Treatment	0.1003***	0.2326***	0.0026	0.0271
	(0.0259)	(0.0838)	(0.0233)	(0.0255)
Number of treated to 1 km	-0.0025	-0.0444***	-0.0073	-0.0078*
	(0.0119)	(0.0150)	(0.0096)	(0.0041)
Observations	888	784	888	784
Mean of dependent variable	0.1196	0.1210	0.1043	0.1048
Panel D: radius of 1.5 km				
Treatment	0.0963***	0.2604***	-0.0058	0.0406*
	(0.0270)	(0.0851)	(0.0244)	(0.0244)
Number of treated to 1.5 km	-0.0052	-0.0030	-0.0012	-0.0024**
	(0.0117)	(0.0027)	(0.0081)	(0.0010)
Observations	877	795	877	795
Mean of dependent variable	0.1209	0.1194	0.1143	0.0928
Number of eligible control	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

[1] Marginal effects of probit model. Clustered standard errors are shown in parenthesis.
[2] Individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[3] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table A.3: Treatment Externalities in the Participation in Organizations in the First Follow Up With Community Fixed Effects.

Dependent variables:	Participation in productive organizations		Participation in community organizations	
	(1)	(2)	(3)	(4)
Panel A: radius of 250-meter				
Treatment	0.1792***	0.1811***	0.0237	0.0198
	(0.0564)	(0.0571)	(0.0194)	(0.0191)
Number of treated to 250 m	-0.0027	0.0113	-0.0014	-0.0008
	(0.0129)	(0.0104)	(0.0059)	(0.0060)
Panel B: radius of 500-meter				
Treatment	0.1798***	0.1838***	0.0238	0.0200
	(0.0561)	(0.0575)	(0.0193)	(0.0192)
Number of treated to 500 m	-0.0254***	-0.0073	-0.0087***	-0.0099**
	(0.0095)	(0.0054)	(0.0029)	(0.0046)
Panel C: radius of 1km				
Treatment	0.1755***	0.1827***	0.0255	0.0184
	(0.0568)	(0.0577)	(0.0202)	(0.0189)
Number of treated to 1 km	-0.0335**	-0.0048	-0.0114*	-0.0086
	(0.0131)	(0.0085)	(0.0066)	(0.0062)
Panel D: radius of 1.5 km				
Treatment	0.1705***	0.1773***	0.0230	0.0186
	(0.0542)	(0.0546)	(0.0194)	(0.0185)
Number of treated to 1.5 km	-0.0358***	-0.0169	-0.0027	-0.0052
	(0.0123)	(0.0123)	(0.0042)	(0.0061)
Observations	1672	1672	1672	1672
Number of eligible control	Yes	Yes	Yes	Yes
Baseline outcome control	No	Yes	No	Yes
Individual controls	No	Yes	No	Yes

[1] Linear probability model estimates. Robust clustered standard errors are shown in parenthesis.
[2] Individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[3] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table A.4: Treatment Externalities in Participation for Individuals Above and Below the Median of the Eligible Distribution.

Dependent variables:	Participation in productive organizations		Participation in community organizations	
	Below the median (1)	Above the median (2)	Below the median (3)	Above the median (4)
Panel A: radius of 250 meters				
Treatment	0.1049***	0.2469***	0.0082	0.0254
	(0.0258)	(0.0830)	(0.0233)	(0.0287)
Number of treated to 250 m	0.0239	-0.0040	-0.0001	-0.0058
	(0.0258)	(0.0138)	(0.0209)	(0.0049)
Observations	917	755	917	755
Mean of dependent variable	0.1099	0.1332	0.1078	0.1005
Panel B: 500 meters				
Treatment	0.1007***	0.2579***	0.0160	0.0159
	(0.0260)	(0.0823)	(0.0231)	(0.0264)
Number of treated to 500 m	0.0026	-0.0258***	-0.0098	-0.0077***
	(0.0187)	(0.0098)	(0.0153)	(0.0027)
Observations	924	748	924	748
Mean of dependent variable	0.1318	0.1045	0.1046	0.1045
Panel C: radius of 1km				
Treatment	0.1003***	0.2326***	0.0026	0.0271
	(0.0259)	(0.0838)	(0.0233)	(0.0255)
Number of treated to 1 km	-0.0025	-0.0444***	-0.0073	-0.0078*
	(0.0119)	(0.0150)	(0.0096)	(0.0041)
Observations	888	784	888	784
Mean of dependent variable	0.1196	0.1210	0.1043	0.1048
Panel D: radius of 1.5 km				
Treatment	0.0963***	0.2604***	-0.0058	0.0406*
	(0.0270)	(0.0851)	(0.0244)	(0.0244)
Number of treated to 1.5 km	-0.0052	-0.0030	-0.0012	-0.0024**
	(0.0117)	(0.0027)	(0.0081)	(0.0010)
Observations	877	795	877	795
Mean of dependent variable	0.1209	0.1194	0.1143	0.0928
Number of eligible control	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[3] Dependent variable is an indicator variable that takes the value of 1 if the individual participates in productive in the first follow up and zero otherwise. Independent variables are the treatment group indicator, that takes the value of one if the individual was assigned to treatment group in baseline and zero otherwise and the number of assigned to treatment neighbors to each survey participant at different arbitrary radius: 250 meters, 500 meters, 1km and 1.5km.
[3] All regressions are controlled by the total number of people to each respective radius, baseline outcome control is each outcome variable in baseline and individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[4] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table A.5: Externalities in Participation using the Square of the Number of Neighbors Assigned to Treatment.

Dependent variables:	Participation in productive organizations			Participation in community organizations		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: radius of 250-meter						
Treatment	0.1783***	0.1783***	0.1760***	0.0291	0.0227	0.0196
	(0.0517)	(0.0516)	(0.0498)	(0.0195)	(0.0176)	(0.0178)
Number of treated to 250 m	-0.0061	-0.0061	-0.0029	-0.0083	-0.0093	-0.0075
	(0.0131)	(0.0131)	(0.0127)	(0.0072)	(0.0068)	(0.0064)
Number of treated to 250m squared	0.0001	0.0001	0.0000	0.0002	0.0002	0.0002
	(0.0003)	(0.0003)	(0.0003)	(0.0001)	(0.0001)	(0.0001)
Panel B: radius of 500-meter						
Treatment	0.1762***	0.1763***	0.1748***	0.0286	0.0217	0.0188
	(0.0499)	(0.0499)	(0.0486)	(0.0183)	(0.0166)	(0.0169)
Number of treated to 500 m	-0.0237**	-0.0237**	-0.0210*	-0.0128***	-0.0132***	-0.0116***
	(0.0119)	(0.0119)	(0.0118)	(0.0044)	(0.0044)	(0.0037)
Number of treated to 500m squared	0.0000	0.0000	-0.0000	0.0001**	0.0001*	0.0001*
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Panel C: radius of 1km						
Treatment	0.1667***	0.1668***	0.1678***	0.0265	0.0198	0.0178
	(0.0470)	(0.0470)	(0.0466)	(0.0188)	(0.0170)	(0.0173)
Number of treated to 1 km	-0.0322**	-0.0322**	-0.0299**	-0.0118**	-0.0113***	-0.0094**
	(0.0142)	(0.0142)	(0.0146)	(0.0046)	(0.0043)	(0.0042)
Number of treated to 1km squared	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Panel D: radius of 1.5 km						
Treatment	0.1604***	0.1604***	0.1616***	0.0267	0.0205	0.0187
	(0.0431)	(0.0431)	(0.0430)	(0.0185)	(0.0170)	(0.0174)
Number of treated to 1.5 km	-0.0345***	-0.0345***	-0.0341***	-0.0065*	-0.0059*	-0.0045
	(0.0132)	(0.0132)	(0.0131)	(0.0036)	(0.0034)	(0.0034)
Number of treated to 1.5km squared	0.0000	0.0000	-0.0000	0.0000	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Observations	1672	1672	1672	1672	1672	1672
Total number of people control	Yes	Yes	Yes	Yes	Yes	Yes
Baseline outcome control	No	Yes	Yes	No	Yes	Yes
Individual controls	No	No	Yes	No	No	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[3] Dependent variable is an indicator variable that takes the value of 1 if the individual participates in productive in the first follow up and zero otherwise. Independent variables are the treatment group indicator, that takes the value of one if the individual was assigned to treatment group in baseline and zero otherwise and the number of assigned to treatment neighbors to each survey participant at different arbitrary radius: 250 meters, 500 meters, 1km and 1.5km.
[3] All regressions are controlled by the total number of people to each respective radius, baseline outcome control is each outcome variable in baseline and individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[4] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table A.6: Incentives to Free-Ride by Neighbors Characteristics (1km Distance).

Dependent variables:	Participation in productive organizations		Participation in community organizations	
	(1) Communal	(2) Private	(3) Communal	(4) Private
Panel A: communal property rights				
Number of treated with same land	-0.0152		-0.0191	
	(0.0249)		(0.0438)	
Number of treated without the same land	-0.0968	-0.0308**	-0.4413	-0.0479**
	(0.0926)	(0.0128)	(0.3187)	(0.0211)
Observations	293	1379	293	1379
	Literate (1)	Illiterate (2)	Literate (3)	Illiterate (4)
Panel B: literacy				
Number literate treated	-0.0109	-0.0388	0.0050	-0.0183
	(0.0183)	(0.0241)	(0.0074)	(0.0120)
Number of illiterate treated	-0.0244	-0.0256	-0.0107	-0.0120
	(0.0190)	(0.0215)	(0.0096)	(0.0116)
Observations	1176	496	1176	496
	Not poor (1)	Poor (2)	Not poor (3)	Poor (4)
Panel C: poverty				
Number of not poor treated	-0.0466**	-0.0266	-0.0023	-0.0276
	(0.0186)	(0.0219)	(0.0080)	(0.0171)
Number of poor treated	-0.0265	0.0039	-0.0140	-0.0051
	(0.0372)	(0.0130)	(0.0199)	(0.0079)
Observations	837	835	837	835
	More productive (1)	Less productive (2)	More productive (3)	Less productive (4)
Panel D: bean productivity				
Number of more productive treated	-0.033**	-0.034***	-0.007	-0.014***
	(0.016)	(0.009)	(0.005)	(0.004)
Number of less productive treated	-0.003	-0.032	0.010	-0.019
	(0.037)	(0.020)	(0.015)	(0.013)
Observations	835	837	835	837
	More productive (1)	Less productive (2)	More productive (3)	Less productive (4)
Panel E: maize productivity				
Number of more productive treated	-0.0866***	-0.0288*	-0.0050	-0.0888***
	(0.0239)	(0.0247)	(0.0161)	(0.0173)
Number of less productive treated	0.0027	-0.0206	0.0041	-0.0104*
	(0.0206)	(0.0143)	(0.0117)	(0.0058)
Observations	509	1163	509	1163
Treatment indicator	Yes	Yes	Yes	Yes
Number of eligible control	Yes	Yes	Yes	Yes
Baseline outcome control	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[2] Individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[2] *** Significant at 1%, ** significant at 5% and * significant at 10%.

Table A.7: Incentives to Free-Ride by Neighbors Characteristics (1.5km Distance).

Dependent variables:	Participation in productive organizations		Participation in community organizations	
	(1) Communal	(2) Private	(3) Communal	(4) Private
Panel A: communal property rights				
Number of treated with same land	0.0413		0.0335**	
	(0.0276)		(0.0113)	
Number of treated without same land	-0.0687	-0.0357***	-0.0128	-0.0040
	(0.0587)	(0.0114)	(0.0289)	(0.0035)
Observations	293	1379	293	1379
	Literate (1)	Illiterate (2)	Literate (3)	Illiterate (4)
Panel B: literacy				
Number literate treated	-0.0264	-0.0119	0.0125	-0.0076
	(0.0174)	(0.0224)	(0.0083)	(0.0113)
Number of illiterate treated	-0.0396**	-0.0258	-0.0069	-0.0090
	(0.0161)	(0.0196)	(0.0083)	(0.0093)
Observations	1176	496	1176	496
	Not poor (1)	Poor (2)	Not poor (3)	Poor (4)
Panel C: poverty				
Number of not poor treated	-0.0403**	-0.0229	0.0002	0.0004
	(0.0201)	(0.0271)	(0.0086)	(0.0107)
Number of poor treated	-0.0643**	-0.0061	-0.0018	-0.0076
	(0.0291)	(0.0129)	(0.0162)	(0.0085)
Observations	837	835	837	835
	More productive (1)	Less productive (2)	More productive (3)	Less productive (4)
Panel D: bean productivity				
Number of more productive treated	-0.038**	-0.034***	-0.010*	0.001
	(0.016)	(0.008)	(0.005)	(0.004)
Number of less productive treated	-0.018	-0.000	0.003	-0.028
	(0.030)	(0.018)	(0.014)	(0.011)
Observations	835	837	835	837
	More productive (1)	Less productive (2)	More productive (3)	Less productive (4)
Panel E: maize productivity				
Number of more productive treated	-0.1068***	-0.0900***	0.0084	-0.0318**
	(0.0232)	(0.0183)	(0.0145)	(0.0132)
Number of less productive treated	-0.0019	-0.0167	0.0018	0.0037
	(0.0146)	(0.0116)	(0.0110)	(0.0053)
Observations	509	1163	509	1163
Treatment indicator	Yes	Yes	Yes	Yes
Number of eligible control	Yes	Yes	Yes	Yes
Baseline outcome control	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes

[1] Marginal effects of probit model. Cluster standard errors at community level are shown in parenthesis.
[2] Individual controls are sex, age, knows to read and write, belongs to an indigenous group, number of household members, agriculture as primary activity, cattle raising as primary activity and corn and bean productivity in apante season.
[2] *** Significant at 1%, ** significant at 5% and * significant at 10%.