From Learning to Doing:
Diffusion of Agricultural Innovations in Guinea-Bissau *

Rute M. Caeiro†

Abstract

This paper analyzes the role of social networks in the diffusion of cultivation techniques introduced by an agricultural project in Guinea-Bissau. We take advantage of this intervention to study the diffusion of knowledge and adoption of cultivation techniques from project participants to the wider community. In order to test for social learning, we exploit a detailed dataset which includes village census and social network data across different social network dimensions. More precisely, we make use of a village photo directory in order to obtain a comprehensive and fully mapped social network dataset. We find evidence that agricultural information spreads across networks from project participants to non-participants. Specific network groups such as the 'regular chatting' and 'agricultural advice' networks appear to be more relevant for that process. Our dataset also allows us to disentangle effects stemming from strong and weak network links, and weak links are found to be equally important in the diffusion of agricultural knowledge. Despite positive diffusion effects in knowledge, we have not found evidence of network effects in adoption behavior.

JEL Classification: O13, Q16, O31.
Keywords: Agriculture, Technology Adoption, Social Networks, Guinea-Bissau, Africa

* I am grateful for helpful suggestions from André Moreira, Inês Vilela, Julia Seither, Paulo Santos and Pedro Vicente. I would like to thank Pedro Santos, Patrícia Maridalho and the VIDA team for their cooperation and support. I wish to acknowledge the financial support provided by European Union and Camões-Instituto da Cooperação e da Língua, through the project "Kopoti pa cudji no futuro". This work was funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209). I would also like to acknowledge FCT for the PhD Scholarship PD/BD/105725/2014.
† Nova School of Business and Economics, Universidade Nova de Lisboa and NOVAFRICA. Email: rute.caeiro2011@novasbe.pt.

1 Introduction

Agricultural development through productivity improvements can be an effective means to reduce poverty amongst Africa's low-income population, for whom agriculture represents the main source of employment and livelihood. Despite some gains in recent years, agricultural productivity in the continent, and in sub-Saharan Africa in particular, remains low, lagging well behind the rest of the world. Although agricultural technologies are available that could significantly boost productivity, those have not yet been widely adopted in the region. Low access to extension services and to reliable information are among the barriers to technology adoption in the agricultural sector, hindering productivity growth. In this context, social interactions may play a key role in mitigating information constraints and disseminating improved technologies.

Agriculture development interventions aiming to increase agricultural productivity and yields have the potential to impact technology adoption beyond project participants, since it is expected that trained farmers will disseminate the techniques to the rest of the community.

This paper analyzes the role of social networks in the diffusion process of improved cultivation techniques introduced by an agricultural project in Guinea-Bissau. The project focused on horticultural production and improved cultivation practices, with the aim of increasing food security and decreasing vulnerability through the diversification of crops and improvements in production practices. We take advantage of this intervention to study the diffusion of improved techniques from project participants to the wider community in one village in Guinea-Bissau. In particular, we test for spillover effects in terms of both knowledge and adoption of production practices. In order to test for social learning effects, we make use of a comprehensive dataset including village census and network data. We collected network data by first asking farmers to list their network contacts. Then, on a second stage, we used a photo directory as a visual aid to help respondents recall additional network links. This network elicitation method allows us to capture more than just the farmers' strong network links, resulting in a more detailed and fully mapped social network. In addition, we collected network data across four network dimensions, allowing us to test for different information channels. Our results suggest that information externalities from the project participants to the rest of the community exist: having a link with a trained farmer is correlated with increases in agricultural knowledge. Furthermore, farmers seem to be mostly learning from their project participant peers, with non-project participant farmers having no effect, or even a marginally significant negative effect on their peer's knowledge. Testing for the different information channels, we find that not all network groups contribute equally for the spread of information. We find stronger network effects for those peers who

farmers regularly chat with or from whom farmers could ask for agricultural advice. In contrast, we find no evidence of information exchange through peers farmers might borrow money from. Furthermore, exploiting our measure of link strength, weak social links appear to be as relevant as strong links in the diffusion of information. Despite positive effects in knowledge, we have not found evidence of social effects on techniques adoption. More specifically, our results indicate that having social connections to project participants has not effect on farmers adoption decisions. These findings are consistent across the different network dimensions.

This paper relates to the literature on diffusion of information, technologies and behavior within social networks. Diffusion effects along social networks have been documented in a variety of settings, including health prevention (Oster and Thornton, 2012; Godlonton and Thornton, 2012; Apouey and Picone, 2014), educational outcomes (Bobonis and Finan, 2009; Fafchamps and Mo, 2017), financial decisions (Duflo and Saez, 2003; Cai, Janvry and Sadoulet, 2015; Banerjee et al., 2013), and agricultural practices (Foster and Rosenzweig, 1995; Munshi, 2004; Bandiera and Rasul, 2006; Conley and Udry, 2010; Van den Broeck and Dercon, 2011).

Our paper is closer in spirit to previous work on agricultural technology diffusion within social networks. Within that, Foster and Rosenzweig, (1995) was one of the early studies to investigate the effects of social learning in agriculture. In a study of adoption of high-yielding varieties in India, the authors found evidence of learning externalities and not just mimicking behavior, with farmers learning from other farmers' experience. More recently, Bandiera and Rasul (2006) studied the adoption of sunflower seeds in Mozambique and provided further evidence of positive peer effects on adoption decisions. In a seminal paper, Conley and Udry (2010), were among the first to study peer-effects on agricultural technology adoption using explicit social network data. Focusing on pineapple producers in Ghana, they identified knowledge flows by asking farmers whether they had gone to a random selection of farmers in the village for agricultural advice. They also found evidence of farmers learning from one another, aligning their use of fertilizers with that of their most successful peers.

Although those and several other studies have found evidence of positive peer effects in technology adoption, unfortunately results have not always been as encouraging. Research has also shown that diffusion of knowledge and practices can be limited in some settings (Fafchamps and Quinn, 2016; Fafchamps and Söderbom, 2014) or non-existent (Duflo, Kremer and Robinson, 2011). And in some contexts, social network effects may even generate perverse outcomes, for example, creating incentives for delaying adoption and free-riding on the experimentation of others (Foster and Rosenzweig, 1995; Bandiera and Rasul, 2006; Maertens, 2017), or making technology adoption less likely altogether (Kremer and Miguel, 2007).

The remainder of this paper is structured as follows. Section 2 provides the context of the study setting. Section 3 describes the setting of the horticultural production project. Section 4 describes the data collection, and the network and outcome measures employed. In Section 5 we outline the estimation strategy. Descriptive statistics are presented in Section 6. Section 7 presents the econometric results. Finally, in Section 8 we conclude.

2 Context

Guinea-Bissau, a country in West Africa with a population of approximately 1.8 million, is one of the poorest countries in the world, with a GDP per capita of 1450$PPP and 67 percent of its population living on less than $1.90 per day.[1] Agriculture is key to Guinea-Bissau's economy: it accounts for 69 percent of GDP and represents the primary source of income for 85 percent of its population. The agriculture sector is dominated by cashew nut production for export, by rice production for consumption, and by horticulture production on a smaller scale.[2] Rice is the main staple crop in the country and is widely grown, but rice productivity has remained relatively low.[3] Low productivity can be explained by the several constraints faced by the agricultural sector, ranging from erratic weather to scarce inputs and extension services, and to weakened infrastructures. Years of poor harvest or shocks to cashew prices can leave subsistence farmers in a particularly vulnerable situation, as was the case in 2012, when a combination of poor cashew harvest, lower export prices and political instability led to a rise in food insecurity.[4] As for the horticultural sector, production has steadily increased, from 26381 tonnes in 2005 to 33420 tonnes in 2014.[5] In the Guinea-Bissau - Country' Economic Memorandum (2015), the World Bank identified horticulture as one of the agricultural sectors with greatest economic potential, and as a potential source of alternative income for rural households that would allow them to mitigate the risks posed by relying on a single cash crop.

The setting for this study was the village of Suzana, in the northwest region of Guinea-Bissau (Figure 1). Suzana is a rural village, with 354 households spread across 8 neighborhoods. The majority of households in Suzana are from the Felupe ethnic group, and most of the individuals in the village are subsistence farmers. As in other regions of the country, rice is the main crop,

[1] World Development Indicators (2017), World Bank.
[2] See World Bank (2015).
[3] In 2014 rice productivity in Guinea-Bissau was 15.6 thousand hectograms per hectare, well below the average African level of 25.9 (FAOSTAT Database).
[4] See World Food Program (2013).
[5] FAOSTAT Database.

and cashew is produced on a smaller scale. Horticultural production is relatively scarce and is almost exclusively a female activity. Furthermore, there are no agriculture extension services in the region.

3 Agricultural intervention

In 2015, the international NGO 'VIDA' introduced an agricultural production project providing agricultural technical training and inputs to farmers, fostering cooperation among them, and promoting linkages to local markets in 6 villages of Guinea-Bissau, including Suzana. The project included training sessions on horticultural cultivation techniques, creation and management of farmers associations, and on the logistics of the supply chain to local markets. This study focuses specifically on the horticultural production component of the project. We take advantage of this intervention in order to study the diffusion effects of cultivation practices from the project participants to the rest of the community in the village of Suzana. In what follows we briefly describe the horticultural production component of the project and the selection of the project participants.

3.1 Horticultural production training

The horticultural production element of the intervention included three modules on cultivation techniques, which took place between November 2015 and February 2016, before the 2016 agricultural season (which runs from March to mid-July), and included a mix of theoretical and practical training sessions focusing mainly on improved production techniques. The training covered practices such as land preparation, irrigation, staking, pruning, soil enrichment, spacing, mulch, seed selection, nursery preparation and management, pest and disease management, organic pesticides, and post-harvesting handling. Although some of those practices were already familiar to farmers, most were newly introduced by the project.

3.2 Project participants

Project participants were selected by the female village leaders, who provided a list of female farmers interested in participating in the intervention. That list of potential participants made up the sample of a randomized impact evaluation conducted on the project. Those were then randomly allocated to either the control or treatment group, with blocks being formed at the level of the village before randomization. In addition, female village leaders also attended the program. Results of the impact evaluation are not the main scope of this paper, but they are briefly addressed

in the next sections. As mentioned in the previous section, this paper focuses only on the village of Suzana, where 35 farmers were assigned to the treatment group and participated in the project, while another 43 constituted the control group. As such those 78 individuals constitute the experimental subjects within Suzana village.

4 Measurement and data

Data collection took place in two stages, by means of two surveys. Survey respondents were the household representatives, chosen to be the individuals responsible for the households' horticulture production - usually the female head of the household. First, a village census was conducted in February and March 2016, including questions on demographic characteristics, horticulture production patterns during the 2015 agricultural season, and household asset ownership. In this first survey, an enumerator took a photo of each respondent in order to compile a photo album of the village. This survey included all 354 households in the village.
The second stage of data collection took place between August and September of 2016, right after the 2016 agricultural season. At this stage we collected data on the individuals' network links, on horticulture production patterns during the 2016 agricultural season, and on horticultural production practices and knowledge. In this survey we were only able to track 345 out the 354 households. Both data collection activities took place after the horticultural training intervention described in the previous section.

4.1 Network measures

The social network was measured by conducting a village census and collecting network data on all the households in the village. This approach therefore allows us to obtain a complete network map. The network data were collected at the second-round survey. We first asked farmers to name all individuals with whom they had a network link in different social scenarios. We did this in an open manner – i.e., not imposing a limit on the number of links the individual could list. However, this method might tend to only capture the individual strong lings, as those closer to the respondent are more likely to be named while leaving out weak links and/or links that the respondent might have forgotten (Maertens and Barrett, 2012, Brewer, 2000). To address this concern, we made use of the village photo album in order to elicit such weak or forgotten links. This elicitation method allowed to obtain a more comprehensive and detailed network map.
The village photo album was organized by neighborhood, included the photo of one person per household (the household representative), and depicted all the households in the village. See Figure 2 for a sample page of the village photo album.

We elicited the individuals' social links along four dimensions: i) kinship network, which is defined as individuals with whom the respondent has a kinship tie, ii) chatting network, which includes individuals the respondent regularly chats with, iii) agricultural advice network, which contains individuals the respondent would go to for agricultural advice, and iv) money-borrowing network, defined as the set of individuals the respondent could ask for money in time of need. The same set of four network questions were repeated in the same order 8 times, one per neighborhood in the village. Furthermore, the order of the neighborhoods was randomized per individual, to prevent bias.

More specifically, the network variables were collected through survey questions in a two-step procedure: 1st Elicit link from "memory", by asking the respondent to list all possible links that he or she might have in that specific neighborhood and without resorting to the album, 2nd Elicit extra links not mentioned yet using the village photo album. Therefore, for the kinship network we first elicited the kinship relationships from "memory" phrasing it as: "Who are your family members that live in the neighborhood of «Catama» but outside of your household residence?". In the second step we elicited the extra links not mentioned yet by asking the respondents to go through the household's representative photos of that neighborhood and asked: "Do you have any other family member living in the neighborhood of «Catama»?". This same procedure was used for the chatting network, which elicited the individuals the respondent talked to on a regular basis (at least once a week) and was phrased as: "Who are the individuals that live in neighborhood of «Catama» with whom you talk on a regular basis (at least once a week)?". This was followed by the agricultural advice network which elicited the individuals that the respondent would go to for agricultural advice, and it was phrased as: "If you had a problem in your plot with one of your crops, to whom would you go for advice in the neighborhood of «Catama»?". Lastly, the fourth network dimension was the borrowing money network which elicited the individuals that the respondent could ask for money in times of need, phrased as: "If you had to borrow money for an unexpected emergency (for example, to buy medication or cloth for a funeral) who would you ask in the neighborhood of «Catama»?". For the latter two network dimensions, namely agriculture advice and borrowing money, we recorded not only the potential link but also the effective link. Hence, after the respondent listed an individual in either the agricultural advice or borrowing money survey questions we would ask if he or she had indeed asked for agricultural advice or borrowed money from that individual in the last 12 months, respectively.

By eliciting social networks as previously described we would be implicitly capturing the strength of ties between individuals, since links elicited from "memory" are more likely to capture strong ties, while the remaining links elicited resorting to the village photo album would more likely

capture weak ties. In what follows, we define strong ties as the links provided from "memory", and weak ties as the links identified using the photo album visual aid. Appendix A provides additional discussion and robustness checks on the measures of tie strength.

Finally, in addition to the aforementioned measurements, we also collected data on group membership, such as religious or self-help groups, as well as neighboring plots through GPS coordinates. In the results that follow we focus on the first four network variables described: kinship, regular chatting, agricultural advice and borrowing money.

4.2 Outcome measures

In this study we focus mainly on two outcome variables of interest: practice adoption, and knowledge. A list of 11 survey questions on production practices, based on the topics covered by the horticultural training, was used to measure the adoption of improved practices. These were then followed by 11 survey questions designed to measure the respondents' knowledge with respect to those same practices. Practices covered included land preparation, irrigation, nursery management, spacing, mulch, soil enrichment, pruning, staking, pest management and crop rotation. The practices adoption questions focused on whether respondents had adopted the aforementioned practices in the previous agricultural season, which had just finished. The practices knowledge questions tested respondents' knowledge on either how to apply or the benefits of those practices. Responses to the two sets of questions were then used to construct two indices, one for production practices adoption and one for production practices knowledge, as the simple average of the z-scores for the relevant survey questions.[6]

5 Estimation strategy

We start our analysis by estimating the treatment effects for the outcomes of interest in the impact evaluation sample. Given the random assignment of the treatment, the average treatment effects of the agriculture training program can be estimated using the specification:

$$Y_i = \alpha + \partial T_i + \varepsilon_i, \quad \textbf{(1)}$$

[6] The z-scores were computed by subtracting the mean and dividing by the standard deviation of the impact evaluation control group. Following Kling, Liebman, and Katz (2007) if an individual has a response to at least one of the 11 survey questions, then any missing value for the other variables are imputed by the group mean.

where Y_i represents the outcome variable of interest for individual i, namely practices knowledge and adoption. T_i is a binary variable which takes the value of one if the individual was assigned to the treatment group and zero otherwise.

The above specification can also be estimated using individual control variables:

$$Y_i = \alpha + \partial T_i + \gamma X_i + \varepsilon_i, \quad \textbf{(2)}$$

where X_i is a vector of individual and household characteristics, such as age, years of education, religion dummies, marital status, whether the households produced horticultural crops in the previous year and household assets.

Average treatment effects are not the primary focus of this paper, however. Instead, we are interested in estimating the diffusion effects of the training program. In order to do so we proceed as follows. First, we test whether any two farmers are more likely to have information about, and to adopt the agricultural practices if they are network partners, and one or both farmers are treated individuals. To this end we consider each individual as a node in the dyadic framework of our data. We use dyadic regressions to test whether farmers are more likely to have higher average outcomes when a link with a treated farmer exists. We take the undirected dyads as the unit of observation, in which individuals are considered linked if either named the other as a network partner, i.e. the direction of the link is disregarded. In what follows we refer to the project participants as treated individuals and to the household representatives of the rest of the community as non-treated. Our specification is:

$$\bar{Y}_{ij} = \alpha + \beta_{TT} L_{ij}^{TT} + \beta_{NT} L_{ij}^{NT} + \gamma_1 w_{ij} + \gamma_2 |z_i - z_j| + \gamma_2 (z_i + z_j) + \gamma P_{ij} + \varepsilon_{ij}, \quad \textbf{(3)}$$

In this case, let $\bar{Y}_{ij}$ represent the average outcome of interest between nodes i and j, when a link between i and j exists. Let the binary variable L_{ij}^{TT} represent the existence of a link between nodes i and j, when both nodes are treated. L_{ij}^{TT} takes the value of one if nodes i and j are treated and zero otherwise. The binary variable L_{ij}^{NT} in turn represents the existence of a link between nodes i and j, when either i or j are treated nodes. Therefore L_{ij}^{NT} takes the value of one if one of the nodes is treated and zero otherwise. w_{ij} is a vector of variables describing the relation between i and j, including whether the respondents have the same religion, belong to the same ethnic group, have the same gender and the geographical distance between them. z_i and z_j captures the individual and household level characteristics of i and j, such as years of education, household

assets, marital status and whether the household produced horticultural crops in the previous year. Given that the dyadic relationship is undirected, $\bar{Y}_{ij}$ is equal to $\bar{Y}_{ji}$. As such regressors must enter the specification in a symmetric form, so that the regressors for observation ij are equal to observation ji. We follow Fafchamps and Gubert (2007) and achieve the symmetry requirement by including the characteristics of the individuals in absolute differences and in sums. This approach allows to control for the effects of the differences in characteristics of the nodes, as well as the combined effect of the characteristics in the outcome of interest. Note that the above specification does not offer a causal analysis of network effects from the treated farmers to the non-treated.

Next, we test whether the existence of a link with a treated individual is correlated with practices knowledge and adoption for the non-treated. We take the directed dyad as the unit of observation, in which the direction of the link is taken into account, i.e. node i is linked to node j only if i reported j as a network partner. Note that since we are considering the direction of the link, a link from node i to node j is not the same as a link from node j to node i. Given that we are interested in estimating the influence of the treated nodes on the non-treated nodes we exclude directed links reported by treated nodes from the analysis. We estimate the following specification:

$$Y_i = \alpha + \beta_T L_{ij}^T + \gamma_1 w_{ij} + \gamma_2(z_i - z_j) + \gamma_2(z_i + z_j) + \varepsilon_{ij}, \quad \textbf{(4)}$$

where, Y_i represents the outcome variable of interest for node i when i is non-treated. L_{ij}^T is a binary variable that captures the existence of a link with a treated node. It takes the value of one if i is non-treated and j is treated, and i is linked with j, and zero if both i and j are non-treated nodes and i is linked with j. w_{ij}, z_i and z_j are defined as before. Since we are considering the directionality of the link from node i to node j, there is no symmetry requirement, and so z_i and z_j enter the equation as simple differences and in sums (Fafchamps and Gubert, 2007).

The above specification was also expanded to analyze the effect of strong and weak links with treated individuals:

$$Y_i = \alpha + \beta_{sT} L_{ij}^{sT} + \beta_{wT} L_{ij}^{wT} + \gamma_1 w_{ij} + \gamma_2(z_i - z_j) + \gamma_2(z_i + z_j) + \varepsilon_{ij}, \quad \textbf{(5)}$$

where, L_{ij}^{sT} and L_{ij}^{wT} are binary variables that capture the existence of a link between a non-treated and a treated node, and the link is characterized as either strong or weak, respectively. In this case,

L_{ij}^{sT} takes the value of one if i is non-treated and j is treated, and i has a strong link with j, and zero otherwise. L_{ij}^{wT} is defined in the same manner but for weak links.

Lastly, we test whether the knowledge and adoption behavior of non-treated individuals is affected by the number of treated individuals in their social networks. We conduct a household level analysis employing the following specification:

$$Y_i = \alpha + \beta_T N_i^T + \beta_{nT} N_i^{nT} + \gamma X_i + \theta \bar{X}_{-i} + \varepsilon_i, \quad \textbf{(6)}$$

where, N_i^T is the number of links with treated individuals, and N_i^{nT} are the number links with non-treated individuals in individual i's social network. In addition, X_i is a vector of individual and household characteristics, which include gender, years of education, marital status, religion, ethnic group, whether the household produced horticultural crops in the previous year, and household assets. $\bar{X}_{-i}$ captures the average individual and household characteristics of individuals i's network members, which allows us to control for the fixed characteristics of the other farmers in the network. $\bar{X}_{-i}$ includes the proportion of female respondents, average years of education, proportion of married respondents, proportion of animists, proportion of respondents from the main ethnic group, proportion of households that produced horticultural crops in the previous year and household assets in individuals i's network. Once again, we expand the above specification in order to distinguish between the effects of strong and weak links:

$$Y_i = \alpha + \beta_{sT} N_i^{sT} + \beta_{wT} N_i^{wT} + \beta_{nT} N_i^{nT} + \gamma X_i + \theta \bar{X}_{-i} + \varepsilon_i, \quad \textbf{(7)}$$

where, N_i^{sT} and N_i^{wT} refer to the number of strong and weak links with treated individuals in i's social network, respectively.

All estimations employ OLS. We estimate robust standard errors in all regressions, except for the estimations in a dyadic framework where, following Cameron, Gelbach, and Miller (2011), we use two-way cluster-robust standard errors, clustered at both i and j.

6 Descriptive statistics

In this section we present descriptive statistics for the sample of respondents from the village of Suzana. The sample includes 35 treated individuals and 319 non-treated individuals. Table 1 presents the descriptive statistics for non-treated respondents and differences relative to the treated. Table 1a is split into basic demographics, religion and ethnicity, and occupation. Table

1b presents the descriptive statistics in terms of network centrality variables, namely out-degree, in-degree and betweenness centrality[7], for the different network dimensions. Figures 3 provide a visual representation of the network links among 54 households from one of the neighborhoods in the village. The networks of kinship, regular chatting, agricultural advice and borrowing money are depicted in Figure 3a, 3b, 3c and 3d, respectively.

Looking at the demographic variables in Table 1a, non-treated respondents on average are approximately 52 years old and have 2 years of education. 87 percent of the non-treated respondents are women, animism is the predominant religion, followed by catholicism, and the majority of individuals (88 percent) belong to the Felupe ethnic group. In terms of occupation, most of the individuals are farmers, and a small minority are vendors.

We now turn our attention towards the networks descriptive statistics in Table 1b. On average the non-treated individuals listed 35 kinship links, 21 chatting links, 5 agricultural advice links and 8 money-borrowing links. On the other hand, treated individuals are younger, more likely to be married, less likely to be widowed, and more likely to be catholic than the rest of the village. They are also more likely to name and to be named as a network link in all four network categories, except for borrowing money, in which they are not more likely to list more links than non-treated individuals. Given the selection procedure of the treated individuals differences are to be expected, since these farmers were chosen by the village leader to attend the training sessions.

< Tables 1 around here >

7 Econometric results

Our main goal is to test whether the agricultural training program had spillover effects from the project participants to the wider community in terms of both knowledge and adoption of production practices. We divide the analysis of econometric results into two parts. First, we present the results of the training program' effects on experimental subjects, including balance tests and the estimated average treatment effect of the intervention within the village of Suzana. Then we move on to our analysis of the network effects of the agriculture training program, using both dyadic and household-level specifications. Furthermore, we make use of our data in order to

[7] Out-degree and in-degree is the number of ties a node sends and the number of ties a node receives, respectively. Betweenness centrality, an indicator of node centrality, is the number of shortest paths from all other nodes that pass through that node.

test for social learning across different network dimensions: kinship, regular chatting, agricultural advice and borrowing money.

7.1 Randomized impact evaluation results

This section describes the empirical findings of the impact evaluation.

7.1.1 Balance

Balance tests between the treatment and control groups of the impact evaluation sample are reported in Tables 2a and 2b. These can be used to assess the comparability between the two groups with respect to demographic characteristics and network variables of centrality for each of the five network dimensions. As expected, given the randomization procedure, we do not find any statistically significant difference between the two groups.

< Tables 2 around here >

7.1.2 Treatment effects

This section presents the results of the impact evaluation training program. Our two main outcomes of interest are the index of production practices knowledge and the index of production practices adoption described in section 4.2. Table 3 displays the estimates of treatment effects for each outcome of interest, under specifications (1) and (2).

< Table 3 around here >

The treatment is estimated to have led to an increase in production practices knowledge of 0.197-0.200 standard deviations, statistically significant at the 10 percent level. There is also a clear positive and statistically significant effect of 0.252-0.254 standard deviations of the treatment on our measure of practices adoption in both specifications. These results suggest that the treatment had the desired effect of increasing knowledge of agricultural practices in the treatment group which translated into an increase in adoption of those same practices.

7.2 Social network effects

We now turn to our analysis of the influence of social networks on farmers' knowledge and adoption of cultivation practices. We begin by employing the undirected dyadic specification (3) to investigate the relationship between two farmers being network partners, when one or both farmers is a treated individual, and the knowledge and adoption of practices of the pair. Our outcomes of interest are the average knowledge and adoption of production practices of the dyad. Results for knowledge and adoption practices are shown in Table 4 and 5, respectively. As mentioned before, for each outcome we present the results for four network variables: kinship, regular chatting, agricultural advice and borrowing money.

< Tables 4 and 5 around here >

Results show a positive and statistically significant correlation between being partnered with a treated individual and the outcomes of the pair in both knowledge and adoption in all network specifications. Unsurprisingly, the coefficients are larger in magnitude when both farmers belong to the treatment group, though they remain positive and statistically significant when one of the farmers is a non-treated individual. Although, as mentioned before, these coefficients should be interpreted with caution since they do not provide us with the causal network effects, they appear to indicate that social learning plays role in the knowledge and adoption decisions of the farmers.

Next, we test how the knowledge and adoption behavior of non-treated individuals is affected by the presence of treated individuals in their social networks. More specifically, we explore whether an existing link with treated individuals is correlated with the outcome of interest for non-treated individuals. As before, we present results for the same four network variables, taking into account the link strength by discriminating between strong and weak links. We do this in a dyadic framework by employing specification (4) and (5). Results for the outcome variables for knowledge and adoption are presented in Tables 6 and 7, respectively.

< Tables 6 and 7 around here >

We start by analyzing the effects on practices knowledge in Table 6. Looking at the table, conditional on a link existing, having a link with a treated individual has a positive and statistically significant effect on practices knowledge, for all network groups. In particular, having a kinship link with a treated individual is associated with a 0.046 standard deviation increase in knowledge, significant at the 5 percent level. This result seems to stem mostly from the weak links, since the coefficient for a strong link with treated is negative and not statistically significant, while the effect of a weak link is positive and statistically significant. As for the regular chatting network, having a link with a treated individual is associated with a 0.083 standard deviation increase in

knowledge. Furthermore, while we see significant results for both strong and weak links in the regular chatting network, the magnitude of the effect is larger for weak links. However, we cannot reject equality of the coefficients. Turning to the agricultural advice network we can see a clear positive and statistically significant effect on knowledge from having a link with a treated individual. These correspond to an increase of 0.084 standard deviations independent of link strength, and a 0.075 increase for a strong link, while weak links are still positive but only marginally significant. Finally, links in the money-borrowing network yield a marginally statistically significant increase in knowledge and no significant results when we take in to account the link strength.

In table 7, we display the results of the same network variables on adoption of production practices. Having a link with a treated individual does not seem to increase the farmer's adoption behavior. The only exception is when we turn our attention to the strength of the link in the chatting network: there is a positive and marginally significant effect of having a link with treated on adoption behavior for strong links and a negative effect for weak links.

These results point to the existence of knowledge externalities in all network categories, but stronger effects seem to exist in the regular chatting and agricultural advice network groups. Furthermore, our results seem to suggest no social network effects on adoption behavior.

Lastly, Tables 8 and 9 present the network effects on knowledge and adoption of practices for the household level analysis, employing specification (6) and (7).

< Tables 8 and 9 around here >

As shown in Table 8, the number of links with treated individuals has a positive and statistically significant effect on knowledge for all network variables, except for the money-borrowing network. When we take into account the link strength, the magnitude and significance of the coefficients differ considerably across network variables. As for the number of links with non-treated individuals, all the specifications yield negative coefficients, although they are not always significant across specifications. Furthermore, the number of links with treated individuals is statistically different from the number of links with non-treated individuals for all network variables. Looking at the kinship network first, an additional kin treatment individual is correlated with an increase in farmers' knowledge of 0.040 standard deviations, significant at the 10 percent level. Consistent with the previous results, when we take into account the effects of strong and weak links, we can see that an additional strong link with treated individuals is not significantly correlated with increases in practices knowledge, while an additional weak link increases

knowledge by 0.079 standard deviations, statistically significant at the 1 percent level. We now turn our attention to the regular chatting network: an additional link with a treated individual is associated with a 0.057 standard deviation increase in knowledge, statistically significant at the 1 percent level. Both strong and weak links seem to be significant and there is no statistically difference between them. Regarding the agricultural advice network, an additional treated individual has a positive and statistically significant effect on practices knowledge of 0.059 standard deviations, significant at the 5 percent level. But in the agricultural advice network only strong links with treated individuals seem to matter for knowledge. Lastly, when we consider the money-borrowing network variables all coefficients are positive, but not statistically significant at conventional levels.

In Table 9 we present the results of network effects on practices adoptions. Looking at the table, the number of links with treated individuals does not seem to have any statistically significant effect on practices adoption. The only exception is when we consider the number of weak links with treated in the regular chatting network; this represents a 0.038 standard deviation decrease in adoption, marginally significant at 10 percent level.

Overall, the results described in this section are consistent with social effects existing in knowledge, although these differ considerably across network dimensions. However, despite the existence of positive externalities in practices knowledge, we have not found evidence of social effects on adoption behavior. We plan to revisit this issue in a future version of this paper upon collecting more data, as the absence of positive adoption results at this stage may simply reflect a pattern of delayed adoption whereby farmers learn about a new technology in one agricultural season and begin to apply it in the next.

6 Concluding remarks

This paper analyzes the role of social networks in the diffusion of cultivation techniques introduced by an agricultural project in Guinea-Bissau. In particular, we study the diffusion of knowledge and adoption of improved techniques from project participants to the rest of the community. To do so, we collected detailed census and network data in the village of Suzana and made use of a network elicitation mechanism that allowed us to obtain a comprehensive network map and a characterization of the strength of network ties. In addition, we elicited network membership across four different network dimensions (kinship, regular chatting, agricultural advice, borrowing money), allowing us to examine the role of each in knowledge and adoption diffusion.

Having established that the agricultural project increased knowledge and adoption of practices in project participants, we went on to investigate the prevalence of spillover effects to the rest of the community. Our results indicate that knowledge externalities do exist, particularly in the network of regular chatting and agricultural advice seeking. However, we do not find evidences of network effects in adoption behavior. Furthermore, using our measures of link strength, we have found that weak social links – which conventional network measurements tend to fail to capture appropriately – appear to be as important as strong links in the dissemination of agricultural knowledge. This result highlights that weak ties can also be valuable sources of new information.

Taken together, the contrast between our results in terms of knowledge and adoption behavior could suggest a pattern of delayed adoption whereby farmers learn about a new technology in one agricultural season and begin to apply it in the next We plan to conduct one more round of data collection in order to answer this question and shed further light on the adoption pattern at play.

References

Apouey, B., and G. Picone. 2014. "Social Interactions and Malaria Preventive Behaviors in Sub-Saharan Africa." *Health Economics* 23(9):994–1012.

Bandiera, O., and I. Rasul. 2006. "Social Networks and Technology Adoption in Northern Mozambique." *The Economic Journal* 116(514):869–902.

Banerjee, A., A.G. Chandrasekhar, E. Duflo, and M.O. Jackson. 2013. "The Diffusion of Microfinance." *Science* 341(6144):1236498–1236498.

Bobonis, G.J., and F. Finan. 2009. "Neighborhood Peer Effects in Secondary School Enrollment Decisions." *Review of Economics and Statistics* 91(4):695–716.

Van den Broeck, K., and S. Dercon. 2011. "Information Flows and Social Externalities in a Tanzanian Banana Growing Village." *Journal of Development Studies* 47(2):231–252.

Cai, J., A. de Janvry, and E. Sadoulet. 2015. "Social Networks and the Decision to Insure." *American Economic Journal: Applied Economics* 7(2):81–108.

Conley, T.G., and C.R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100(1):35–69.

Duflo, E., M. Kremer, and J. Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review* 101(6):2350–2390.

Duflo, E., and E. Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 118(3):815–842.

Fafchamps, M., and F. Gubert. 2007. "Risk Sharing and Network Formation." *American Economic Review* 97:75–79.

Fafchamps, M., and D. Mo. 2017. "Peer effects in computer assisted learning: evidence from a randomized experiment." *Experimental Economics*:1–28.

Fafchamps, M., and S. Quinn. 2016. "Networks and Manufacturing Firms in Africa: Results from a Randomized Field Experiment." *The World Bank Economic Review*:lhw057.

Fafchamps, M., and M. Söderbom. 2014. "Network Proximity and Business Practices in African Manufacturing." *The World Bank Economic Review* 28(1):99–129.

Foster, A.D., and M.R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103(6):1176–1209.

Godlonton, S., and R. Thornton. 2012. "Peer effects in learning HIV results." *Journal of Development Economics* 97(1):118–129.

Kremer, M., and E. Miguel. 2007. "The Illusion of Sustainability." *The Quarterly Journal of Economics* 122(3):1007–1065.

Maertens, A. 2017. “Who Cares What Others Think (or Do)? Social Learning and Social Pressures in Cotton Farming in India.” *American Journal of Agricultural Economics* 99(4):988–1007.

Munshi, K. 2004. “Social Learning in a Heterogeneous Population:Technology Diffusion in the Indian Green Revolution.” *Journal of Development Economics* 73(1):185–213.

Oster, E., and R. Thornton. 2012. “Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-Up.” *Journal of the European Economic Association* 10(6):1263–1293.

Table 1a: Individual characteristics - differences across treatment and non-treatment groups

		non-treatment	treatment
basic demographics	**age**	51.652	-9.258***
			(2.128)
	female	0.875	0.095***
			(0.034)
	years of education	1.969	0.266
			(0.522)
	married	0.523	0.271***
			(0.075)
	separated	0.044	0.015
			(0.042)
	single	0.109	0.009
			(0.058)
	widowed	0.324	-0.295***
			(0.039)
religion and ethnicity	**no religion**	0.047	-0.017
			(0.031)
	catholic	0.255	0.186**
			(0.089)
	other christian	0.025	-0.025***
			(0.009)
	animist	0.632	-0.132
			(0.090)
	muslim	0.040	-0.011
			0.031
	felupe	0.882	0.089***
			(0.034)
	balanta	0.040	-0.040***
			(0.011)
	mandinga	0.009	0.020
			(0.030)
occupation	**farmer**	0.749	0.104
			(0.066)
	stays at home	0.167	-0.020
			(0.065)
	vendor	0.019	-0.019**
			(0.008)

Note: Standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1b: Individual characteristics - differences across treatment and non-treatment groups

		non-treatment	treatment
kinship network	**out-degree**	35.184	10.228**
			(4.006)
	in-degree	35.084	11.269**
			(4.463)
	betweenness	0.003	0.001*
			(0.001)
regular chatting network	**out-degree**	20.885	7.674*
			(4.115)
	in-degree	20.685	9.756***
			(2.751)
	betweenness	0.004	0.002**
			(0.001)
agricultural advice network	**out-degree**	5.243	4.728***
			(1.680)
	in-degree	4.919	8.110***
			(1.520)
	betweenness	0.003	0.005***
			(0.002)
borrowing money network	**out-degree**	7.542	0.399
			(1.187)
	in-degree	7.221	3.749**
			(1.542)
	betweenness	0.005	0.003*
			(0.001)

Note: Standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2a: Individual characteristics - differences across treatment and control groups

		control	treatment
basic demographics	**age**	39.439	2.955
			(2.559)
	female	1.000	-0.029
			(0.029)
	years of education	2.000	0.235
			(0.633)
	married	0.780	0.014
			(0.096)
	separated	0.000	0.059
			(0.041)
	single	0.098	0.020
			(0.073)
	widowed	0.122	-0.093
			(0.063)
religion and ethnicity	**no religion**	0.073	-0.044
			(0.051)
	catholic	0.293	0.148
			(0.112)
	other christian	0.024	-0.024
			(0.027)
	animist	0.585	-0.085
			(0.117)
	muslim	0.024	0.005
			0.038
	felupe	0.951	0.019
			(0.046)
	balanta	0.024	-0.024
			(0.027)
	mandinga	0.024	0.005
			(0.038)
occupation	**farmer**	0.902	-0.049
			(0.076)
	stays at home	0.073	0.074
			(0.072)
	vendor	0.024	-0.024
			(0.027)

Note: Standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2b: Individual characteristics - differences across treatment and control groups

		control	treatment
kinship network	**out-degree**	44.634	0.778
			(5.787)
	in-degree	42.976	3.377
			(5.372)
	betweenness	0.004	0.000
			(0.001)
regular chatting network	**out-degree**	27.195	1.364
			(5.132)
	in-degree	29.293	1.148
			(3.300)
	betweenness	0.005	0.001
			(0.001)
agricultural advice network	**out-degree**	10.537	-0.566
			(2.164)
	in-degree	10.220	2.810
			(1.765)
	betweenness	0.006	0.002
			(0.002)
borrowing money network	**out-degree**	10.098	-2.156
			(1.619)
	in-degree	10.000	0.971
			(1.747)
	betweenness	0.007	0.000
			(0.002)

Note: Standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Adoption and knowledge of production practices

dependent variable ------>		practices knowledge		practices adoption	
		(1)	(2)	(3)	(4)
Treatment	coefficient	0.200*	0.197*	0.254***	0.252***
	standard error	(0.116)	(0.111)	(0.095)	(0.096)
mean dep. variable (control)		0.000	0.000	0.000	0.000
r-squared adjusted		0.022	-0.001	0.069	0.053
number of observations		75	75	75	75
controls		no	yes	no	yes

Note: All regressions are OLS. The unit of observation is the individual. Only observations from the impact evaluation sample are included. The dependent variables are an average of z-scores. 'treatment' is a binary variable, which takes the value of one if the individual was assigned to the treatment group and zero otherwise. Controls are individual and household characteristics, which include age, years of education, religion dummies, marital status, whether the households produced horticultural crops in the previous year and household assets. Robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Knowledge of production practices

dependent variable ------>		practices knowledge			
network variable ------>		kinship	regular chatting	agricultural advice	borrowing money
		(1)	(2)	(3)	(4)
undirected link between both treated (β_{TT})	coefficient	0.228***	0.227***	0.267***	0.213**
	standard error	(0.076)	(0.072)	(0.082)	(0.101)
undirected link between non-treated and treated (β_{NT})	coefficient	0.126***	0.128***	0.094**	0.101**
	standard error	(0.033)	(0.040)	(0.037)	(0.045)
mean dep. Variable		-1.704	-1.498	-1.271	-1.908
$\beta_{TT} = \beta_{NT}$	F-stat p-value	0.030	0.004	0.006	0.136
r-squared adjusted		0.409	0.392	0.301	0.447
number of observations		5 888	3 770	1 011	1 200
controls		yes	yes	yes	yes

Note: All regressions are OLS. The unit of observation is the undirected dyad. All undirected dyads are taken as observations. The dependent variable is an average of z-scores. 'undirected link between both treated' and 'undirected link between non-treated and treated' are dummy variables that capture the existence of an undirected link between nodes i and j. 'undirected link between non-treated and treated' takes the value of one if both nodes are treated and zero otherwise. 'undirected link between non-treated and treated' takes the value of one if either i or j. are treated and zero otherwise. Controls include characteristics of the dyad and of both nodes. Dyad controls include whether the respondents have the same religion, belong to the same ethnic group, have the same gender and the geographical distance between them. Node controls are individual and household characteristics, which include years of education, household assets, marital status and whether the household produced horticultural crops in the previous year. Two-way cluster-robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Adoption of production practices

dependent variable ------>		practices adoption			
network variable ------>		kinship	regular chatting	agricultural advice	borrowing money
		(1)	(2)	(3)	(4)
undirected link between both treated (β_{TT})	coefficient	0.209***	0.341***	0.303***	0.194**
	standard error	(0.057)	(0.097)	(0.097)	(0.100)
undirected link between non-treated and treated (β_{TT})	coefficient	0.106***	0.158***	0.131**	0.112**
	standard error	(0.026)	(0.042)	(0.039)	(0.041)
mean dep. Variable		-1.825	-1.781	-1.741	-1.886
$\beta_{TT} = \beta_{NT}$	F-stat p-value	0.003	0.002	0.011	0.277
r-squared adjusted		0.633	0.576	0.488	0.643
number of observations		5 947	3 815	1 016	1 216
controls		yes	yes	yes	yes

Note: All regressions are OLS. The unit of observation is the undirected dyad. All undirected dyads are taken as observations. The dependent variable is an average of z-scores. 'undirected link between both treated' and 'undirected link between non-treated and treated' are dummy variables that capture the existence of an undirected link between nodes i and j. 'undirected link between non-treated and treated' takes the value of one if both nodes are treated and zero otherwise. 'undirected link between non-treated and treated' takes the value of one if either i or j. are treated and zero otherwise. Controls include characteristics of the dyad and of both nodes. Dyad controls include whether the respondents have the same religion, belong to the same ethnic group, have the same gender and the geographical distance between them. Node controls are individual and household characteristics, which include years of education, household assets, marital status and whether the household produced horticultural crops in the previous year. Two-way cluster-robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Knowledge of production practices

dependent variable ------>		practices knowledge							
network variable ------>		kinship		regular chatting		agricultural advice		borrowing money	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
link with treated (β_T)	Coeficiente	0.046**		0.083***		0.084**		0.089*	
	standard error	(0.019)		(0.026)		(0.035)		(0.052)	
strong link with treated (β_{sT})	Coeficiente		-0.008		0.069**		0.075**		0.081
	standard error		(0.029)		(0.035)		(0.037)		(0.053)
weak link with treated (β_{wT})	Coeficiente		0.100***		0.110***		0.115*		0.117
	standard error		(0.030)		(0.037)		(0.067)		(0.080)
mean dep. variable		-1.651	-1.648	-1.528	-1.528	-0.937	-0.935	-1.533	-1.533
$\beta_{sT} = \beta_{wT}$	F-stat p-value		0.020		0.425		0.553		0.609
r-squared adjusted		0.338	0.339	0.313	0.313	0.194	0.194	0.311	0.311
number of observations		11166	11 166	6 669	6 669	1 680	1 680	2 406	2 406
controls		yes	yes	yes	yes	yes	yes	yes	yes

Note: All regressions are OLS. The unit of observation is the directed dyad. Observations with links reported by treated nodes are not included. The dependent variable is an average of z-scores. 'link with treated', 'strong link with treated' and 'weak link with treated' are dummy variables that captures the existence of a link between nodes. 'link with treated' takes the value of one if i is non-treated and j is treated, and zero if both i and j are non-treated nodes. 'strong link with treated' takes the value of one if i is non-treated and j is treated, and the link is characterized as strong, and zero otherwise. 'weak link with treated' takes the value of one if i is non-treated and j is treated, and the link is characterized as weak, and zero otherwise. Controls include characteristics of the dyad and of both nodes. Dyad controls include whether the respondents have the same religion, belong to the same ethnic group, have the same gender and the geographical distance between them. Node controls are individual and household characteristics, which include years of education, household assets, marital status and whether the household produced horticultural crops in the previous year. Two-way cluster-robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Adoption of production practices

dependent variable ------>		practices adoption							
network variable ------>		kinship		regular chatting		agricultural advice		borrowing money	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
link with treated (β_{TT})	Coeficiente	0.014		0.005		0.063		0.041	
	standard error	(0.013)		(0.021)		(0.053)		(0.038)	
strong link with treated (β_{TT})	Coeficiente		0.006		0.050*		0.077		0.047
	standard error		(0.017)		(0.028)		(0.052)		(0.046)
weak link with treated (β_{TT})	Coeficiente		0.022		-0.083*		0.016		0.022
	standard error		(0.021)		(0.045)		(0.072)		(0.070)
mean dep. variable		-1.895	-1.894	-2.072	-2.069	-1.992	-1.996	-1.949	-1.949
$\beta_{TT} = \beta_{NT}$	F-stat p-value		0.570		0.027		0.232		0.776
r-squared adjusted		0.582	0.582	0.488	0.489	0.428	0.428	0.567	0.567
number of observations		11 217	11 217	6 697	6 697	1 680	1 680	2 420	2 420
Controls		yes	yes	yes	yes	yes	yes	yes	yes

Note: All regressions are OLS. The unit of observation is the directed dyad. Observations with links reported by treated nodes are not included. The dependent variable is an average of z-scores. 'link with treated', 'strong link with treated' and 'weak link with treated' are dummy variables that captures the existence of a link between nodes. 'link with treated' takes the value of one if i is non-treated and j is treated, and zero if both i and j are non-treated nodes. 'strong link with treated' takes the value of one if i is non-treated and j is treated, and the link is characterized as strong, and zero otherwise. 'weak link with treated' takes the value of one if i is non-treated and j is treated, and the link is characterized as weak, and zero otherwise. Controls include characteristics of the dyad and of both nodes. Dyad controls include whether the respondents have the same religion, belong to the same ethnic group, have the same gender and the geographical distance between them. Node controls are individual and household characteristics, which include years of education, household assets, marital status and whether the household produced horticultural crops in the previous year. Two-way cluster-robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: Knowledge of production practices

dependent variable ------>		practices knowledge							
network variable ------>		kinship		regular chatting		agricultural advice		borrowing money	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
number of links with treated (β_T)	Coeficiente	0.040*		0.057***		0.059**		0.059	
	standard error	(0.020)		(0.022)		(0.024)		(0.039)	
number of strong links with treated (β_{sT})	Coeficiente		-0.009		0.052*		0.070**		0.049
	standard error		(0.029)		(0.031)		(0.027)		(0.044)
number of weak links with treated (β_{wT})	Coeficiente		0.079***		0.062**		0.029		0.094
	standard error		(0.025)		(0.024)		(0.054)		(0.073)
number of links with non-treated (β_{nT})	Coeficiente	-0.004	-0.004	-0.009*	-0.009*	-0.014	-0.015*	-0.018*	-0.019*
	standard error	(0.004)	(0.004)	(0.005)	(0.005)	(0.009)	(0.009)	(0.011)	(0.011)
mean dep. Variable		-0.910	-0.930	-0.664	-0.656	-1.046	-1.042	-1.083	-1.094
$\beta_T = \beta_{nT}$	F-stat p-value	0.063		0.011		0.017		0.091	
$\beta_{sT} = \beta_{wT}$	F-stat p-value		0.021		0.770		0.512		0.590
r-squared adjusted		0.334	0.344	0.354	0.352	0.456	0.454	0.337	0.335
number of observations		308	308	308	308	308	308	308	308
Controls		yes	yes	yes	yes	yes	yes	yes	yes

Note: All regressions are OLS. The unit of observation is the household. Treated households are excluded from the observations. The dependent variable is an average of z-scores. 'number of links with treated' is the number of links with treated individuals in individual i's social network. 'number of links with non-treated' is the number of links with non-treated individuals in individual i's social network. 'number of strong links with treated' and 'number of weak links with treated' refer to the number of strong and weak links with treated individuals in i's social network, respectively. Controls are demographic characteristics and average demographic characteristics in the network. Demographic characteristics include gender, years of education, marital status, religion, ethnic group, whether the household produced horticultural crops in the previous year, and household assets. Average demographic characteristics in the network include proportion of female respondents, average years of education, proportion of married respondents, proportion of animists, proportion of respondents from the main ethnic group, proportion of households that produced horticultural crops in the previous year and household assets in the network. Robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9: Adoption of production practices

dependent variable ------>		practices adoption							
network variable ------>		Kinship		regular chatting		agricultural advice		borrowing money	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
number of links with treated (β_T)	coefficient	0.017		-0.008		0.033		0.022	
	standard error	(0.016)		(0.018)		(0.034)		(0.027)	
number of strong links with treated (β_{sT})	coefficient		0.009		0.024		0.051		0.023
	standard error		(0.022)		(0.026)		(0.043)		(0.029)
number of weak links with treated (β_{wT})	coefficient		0.023		-0.038*		-0.019		0.018
	standard error		(0.020)		(0.022)		(0.051)		(0.059)
number of links with non-treated (β_{nT})	coefficient	0.000	0.000	-0.001	-0.001	0.004	0.003	-0.000	0.000
	standard error	(0.003)	(0.003)	(0.004)	(0.004)	(0.008)	(0.009)	(0.008)	(0.008)
mean dep. Variable		-2.284	-2.288	-1.282	-1.329	-1.824	-1.817	-1.908	-1.907
$\beta_T = \beta_{nT}$	F-stat p-value	0.373		0.761		0.467		0.491	
$\beta_{sT} = \beta_{wT}$	F-stat p-value		0.592		0.048		0.275		0.931
r-squared adjusted		0.574	0.573	0.562	0.567	0.626	0.627	0.567	0.565
number of observations		311	311	311	311	311	311	311	311
Controls		yes	yes	yes	yes	yes	yes	yes	yes

Note: All regressions are OLS. The unit of observation is the household. Treated households are excluded from the observations. The dependent variable is an average of z-scores. 'number of links with treated' is the number of links with treated individuals in individual i's social network. 'number of links with non-treated' is the number of links with non-treated individuals in individual i's social network. 'number of strong links with treated' and 'number of weak links with treated' refer to the number of strong and weak links with treated individuals in i's social network, respectively. Controls are demographic characteristics and average demographic characteristics in the network. Demographic characteristics include gender, years of education, marital status, religion, ethnic group, whether the household produced horticultural crops in the previous year, and household assets. Average demographic characteristics in the network include proportion of female respondents, average years of education, proportion of married respondents, proportion of animists, proportion of respondents from the main ethnic group, proportion of households that produced horticultural crops in the previous year and household assets in the network. Robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.



Figure 1: Map of Guinea Bissau



Figure 2: Village photo album sample page

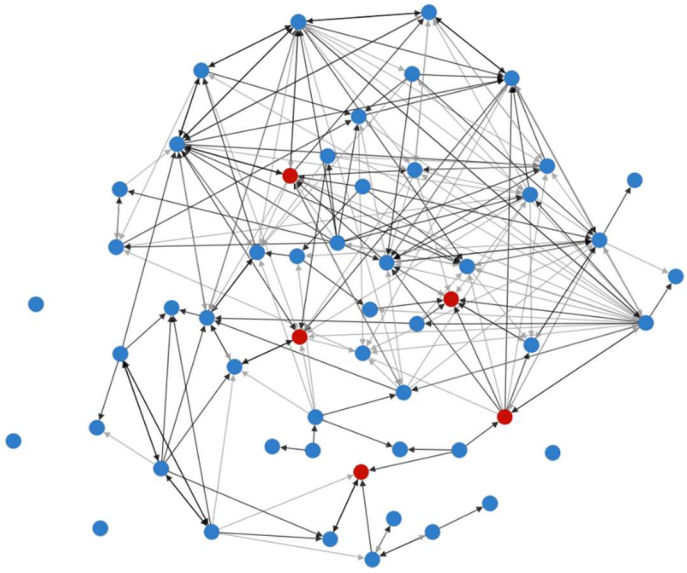


Figure 3a: Illustration of a kinship network

Note: Visual representation of the kinship links among 54 households from one of the neighborhoods in the village. Each node represents a household. Treated households are depicted as red nodes, while non-treated households are represented as blue nodes. The lines between nodes indicate the existence of a link and the direction of the link is illustrated by the arrow. Grey lines represent weak links and black lines represent strong links.

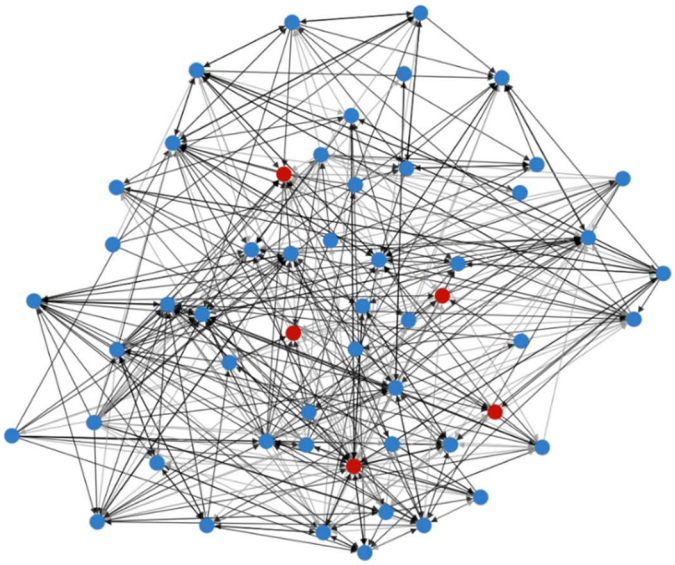


Figure 3b: Illustration of a regular chatting network

Note: Visual representation of the regular chatting links among 54 households from one of the neighborhoods in the village. Each node represents a household. Treated households are depicted as red nodes, while non-treated households are represented as blue nodes. The lines between nodes indicate the existence of a link and the direction of the link is illustrated by the arrow. Grey lines represent weak links and black lines represent strong links.

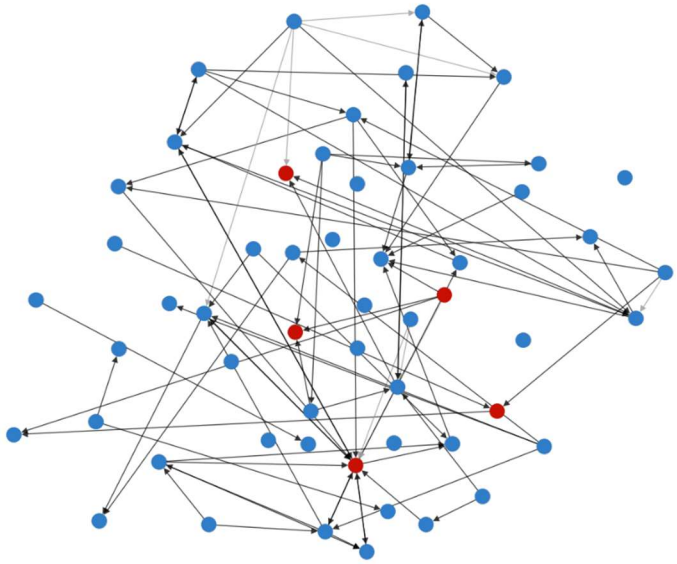


Figure 3c: Illustration of an agricultural advice network

Note: Visual representation of the agricultural advice links among 54 households from one of the neighborhoods in the village. Each node represents a household. Treated households are depicted as red nodes, while non-treated households are represented as blue nodes. The lines between nodes indicate the existence of a link and the direction of the link is illustrated by the arrow. Grey lines represent weak links and black lines represent strong links.

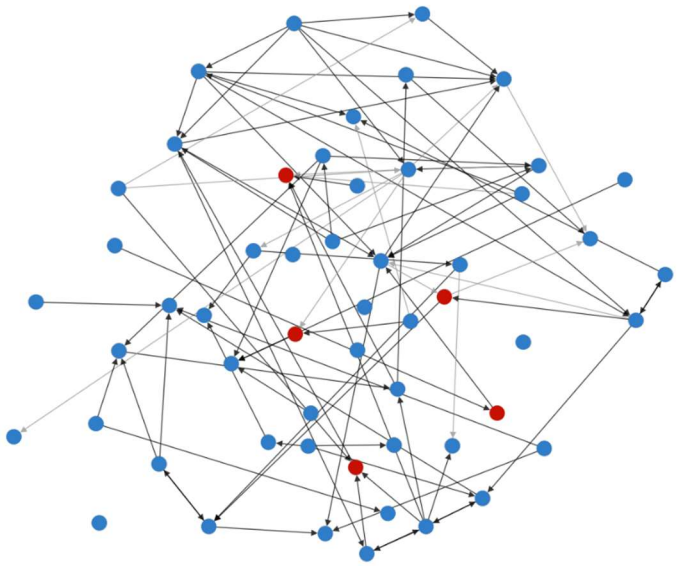


Figure 3d: Illustration of a borrowing money network

Note: Visual representation of the borrowing money links among 54 households from one of the neighborhoods in the village. Each node represents a household. Treated households are depicted as red nodes, while non-treated households are represented as blue nodes. The lines between nodes indicate the existence of a link and the direction of the link is illustrated by the arrow. Grey lines represent weak links and black lines represent strong links.

Appendix A

This appendix expands on our measure of tie strength presented in Section 3.1. As mentioned in the main text, recall-based elicitation methods of collecting network data might result in only capturing the individual's strongest ties. Given our elicitation method, we believe that the links elicited from memory would tend to capture stronger ties, while further links elicited with the album visual aid would more likely represent weak ones. According to Granovetter, 1973, in his seminal paper, "the strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie". In practice, different proxies have been used in order to characterize tie strength, such as reciprocity of the link and the number of mutual friends (Gee, Jones and Burke, 2017). As a robustness check, we test the relationship between our measure of tie strength and some of those proxies.

We follow a dyadic approach, using the directed dyad as the unit of observation. In this case the dyad is a pair of linked nodes and the directionality of the link is taken into account.[1] We estimate the following specification in a dyadic framework:

$$Y_{ij} = \alpha + \beta L_{ij}^{S} + \gamma_1 w_{ij} + \gamma_2(z_i - z_j) + \gamma_2(z_i + z_j) + \varepsilon_{ij}, \quad \textbf{(1)}$$

where Y_{ij} are the proxies for tie strength between nodes i and j: link reciprocity and the proportion of mutual ties. Link reciprocity is a binary variable, taking the value of one if there is a reciprocal relationship between nodes i and j, i.e. if both named the other as a network partner., and of zero if the relationship is unilateral, i.e. if node i named node j as a network partner, but not the other way around. The proportion of mutual ties of nodes i and j is the number of network partners common to i and j divided by the total number of network partners of both i and j. L_{ij}^{S} is a binary network variable that captures tie strength for directed links. It takes the value of one if the link was elicited from memory (strong link), and of zero if it was elicited with the album visual aid (weak link). w_{ij} is a vector of variables describing the characteristics of the dyad, including whether nodes i and j have the same religion, belong to the same ethnic group, are of the same gender, and the geographical distance between them. z_i and z_j are vectors of individual and household level characteristics of i and j, such as years of education, household assets, marital status, and whether the household produced horticultural crops in the previous year. We follow Fafchamps and Gubert (2007) and include the characteristics of the individual and household-

[1] Household i is linked to household j, if household i named household j as a network partner.

level characteristics as simple differences and in sums. By including the regressors in this manner we are able to account for the effects of the differences in characteristics of the nodes, as well as the combined effect of those characteristics. All estimations are OLS and we use two-way cluster-robust standard errors, clustered at both i and j, following Cameron, Gelbach, and Miller (2011).

We present the results for the aforementioned specifications in Tables A1. Table A1a, A1b, A1c and A1d report the results for the network of kinship, regular chatting, agricultural advice and borrowing money, respectively.

< Tables A1 around here >

As we can see from Table A1a having a strong kinship link in our measure is associated with a 3.7 percentage point increase in link reciprocity, and a 0.028 increase in the proportion of mutual ties. Both results are statistically significant at the 1 percent level. As for the network of regular chatting in Table A1b, a strong regular chatting link has a positive and statistically significant correlation with link reciprocity and mutual ties. These represent a 10 percentage points increase on link reciprocity and a 0.041 increase in the proportion of mutual ties. Regarding the network of agricultural advice in Table A1c, we see similar results in link reciprocity: a strong agricultural advice link increases the probability of the link being reciprocal by 4.5 percentage points, statistically significant at the 1 percent level. However, there is no statistically significant effect on the proportion of mutual ties. Lastly, in Table A1d we display the results for the borrowing money network. In line with the results found before, a strong borrowing money link is associated with a 3.9 percentage point increase in link reciprocity and a 0.034 increase in the proportion of mutual ties between the nodes.

Overall, there is a positive correlation between strong links and link reciprocity for all network variables. Similar results arise using the proportion of mutual ties instead of reciprocity: having a strong link in any network category is generally associated with a higher proportion of mutual ties relative to weak links. The sole exception are agricultural advice links, for which coefficients are not significant. These results support our network definition of tie strength, i.e. that links recalled from memory are more likely to be strong than links recalled using the visual aid.

Table A1a: Kinship link strength

dependent variable ------>		link reciprocity	mutual ties
		(1)	(2)
strong kinship link	coefficient	0.037***	0.028***
	standard error	(0.012)	(0.004)
mean dep. Variable		0.011	0.276
r-squared adjusted		0.021	0.055
number of observations		12 571	12 571
controls		yes	yes

Note: All regressions are OLS. The unit of observation is the directed dyad. The dependent variables link reciprocity is binary. It takes the value of one if there is a reciprocal relationship between nodes i and j, and zero otherwise. The dependent variable proportion of mutual ties is the number ties common to nodes i and j divided by the total number of ties in both i and j. 'strong kinship link' is a dummy variable. It takes the value of one if nodes i and j have a strong kinship link and zero if the link is weak. Controls include characteristics of the dyad and of both nodes. Dyad controls include whether the respondents have the same religion, belong to the same ethnic group, have the same gender and the geographical distance between them. Node controls are individual and household characteristics, which include years of education, household assets, marital status and whether the household produced horticultural crops in the previous year. Two-way cluster-robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A1b: Regular chatting link strength

dependent variable ------>		link reciprocity	mutual ties
		(1)	(2)
strong regular chatting link	coefficient	0.100***	0.041***
	standard error	(0.011)	(0.005)
mean dep. Variable		-0.118	0.226
r-squared adjusted		0.040	0.105
number of observations		7 604	7 604
controls		yes	yes

Note: All regressions are OLS. The unit of observation is the directed dyad. The dependent variables link reciprocity is binary. It takes the value of one if there is a reciprocal relationship between nodes i and j, and zero otherwise. The dependent variable proportion of mutual ties is the number ties common to nodes i and j divided by the total number of ties in both i and j. 'strong regular chatting link' is a dummy variable. It takes the value of one if nodes i and j have a strong regular chatting link and zero if the link is weak. Controls include characteristics of the dyad and of both nodes. Dyad controls include whether the respondents have the same religion, belong to the same ethnic group, have the same gender and the geographical distance between them. Node controls are individual and household characteristics, which include years of education, household assets, marital status and whether the household produced horticultural crops in the previous year. Two-way cluster-robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A1c: Agricultural advice link strength

dependent variable ------>		link reciprocity	mutual ties
		(1)	(2)
strong agricultural advice link	coefficient	0.045***	0.006
	standard error	(0.016)	(0.011)
mean dep. Variable		-0.195	0.576
r-squared adjusted		0.029	0.129
number of observations		2 010	2 010
controls		yes	yes

Note: All regressions are OLS. The unit of observation is the directed dyad. The dependent variables link reciprocity is binary. It takes the value of one if there is a reciprocal relationship between nodes i and j, and zero otherwise. The dependent variable proportion of mutual ties is the number ties common to nodes i and j divided by the total number of ties in both i and j. 'strong agricultural advice link' is a dummy variable. It takes the value of one if nodes i and j have a strong agricultural advice link and zero if the link is weak. Controls include characteristics of the dyad and of both nodes. Dyad controls include whether the respondents have the same religion, belong to the same ethnic group, have the same gender and the geographical distance between them. Node controls are individual and household characteristics, which include years of education, household assets, marital status and whether the household produced horticultural crops in the previous year. Two-way cluster-robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A1d: Borrowing money link strength

dependent variable ------>		link reciprocity	mutual ties
		(1)	(2)
strong borrowing money link	coefficient	0.039***	0.034***
	standard error	(0.015)	(0.011)
mean dep. Variable		-0.042	0.466
r-squared adjusted		0.030	0.132
number of observations		2 665	2 665
controls		yes	yes

Note: All regressions are OLS. The unit of observation is the directed dyad. The dependent variables link reciprocity is binary. It takes the value of one if there is a reciprocal relationship between nodes i and j, and zero otherwise. The dependent variable proportion of mutual ties is the number ties common to nodes i and j divided by the total number of ties in both i and j. 'strong borrowing money link' is a dummy variable. It takes the value of one if nodes i and j have a strong borrowing money link and zero if the link is weak. Controls include characteristics of the dyad and of both nodes. Dyad controls include whether the respondents have the same religion, belong to the same ethnic group, have the same gender and the geographical distance between them. Node controls are individual and household characteristics, which include years of education, household assets, marital status and whether the household produced horticultural crops in the previous year. Two-way cluster-robust standard errors reported in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%.