Social Networks and Technology Adoption

Evidence from a Randomised Controlled Trial in Kenya

Varun Satish

Supervised by Prof. Shyamal Chowdhury and Dr. Yi Sun

Abstract

We investigate the relationship between social networks and the adoption of an agricultural technology in Central Kenya. Using social network data collected as part of a Randomised Controlled Trial (RCT), we find that untreated farmers are 2.4 percentage points more likely to adopt the technology if they know an additional treated farmer. We provide evidence that the most important social connections for untreated farmers are ones with households who adopt the technology and not necessarily those which are treated. Moreover, we find that social connections with adopters are just as significant for the adoption outcomes of of untreated farmers in both treatment and control villages. These findings provide evidence that the key mechanism underlying the effect of social connections on adoption outcomes are connections with adopters of the technology.

November 2019

School of Economics
University of Sydney

Thesis submitted in partial fulfilment of the award course requirements of the Bachelor of Economics (Honours)

STATEMENT OF ORIGINALITY

I hereby declare that this submission is my own work and to the best of my knowledge it contains no material previously published or written by another person. Nor does it contain any material which has been accepted for the award of any other degree or diploma at the University of Sydney or at any other educational institution, except where due acknowledgement is made in this thesis.

Any contributions made to the research by others with whom I have had the benefit of working at the University of Sydney is explicitly acknowledged.

I also declare that the intellectual content of this study is the product of my own work and research, except to the extent that assistance from others in the project's conception and design is acknowledged.

Varun Satish

November 2019

ACKNOWLEDGEMENTS

I would first like to express my gratitude for my advisors Shyamal Chowdhury and Yi Sun. Your support and guidance has been much appreciated and has not only ensured my sanity, but has made the year an extremely enjoyable one.

To the 2019 Economics Honours cohort, I do not believe I have enough space on this page to express how thankful I am for having been surrounded by such supportive and talented colleagues. The Honours year comes and goes in the blink of an eye, I hope however, that the friendships we have forged will last much longer. In particular I would like to thank Patrick Meyer, Tim Grey and Lea Jurkovic for their eagerness to listen to my problems.

I would also like to thank those from outside the University. My mother, for her years of hard work and sacrifice. My housemates Will, Zoe, Chris and Henry for their friendship and support and most importantly the two dogs - Busy and Lexi, for showering me with their unconditional love.

Finally, I would like to express my gratitude to all of the participants and administrators of the survey. If it were not for the time, energy and passion of individuals who are actively trying to end poverty, we would never have had the privilege of conducting this research.

Varun Satish

November 2019

Contents

1 Introduction

Food production in Kenya is a serious concern. One report has suggested that annual agricultural production would need to increase by an estimated 75% from 2015 levels in order to meet consumption in 2030 (Cilliers et al., 2018). The diffusion of more productive agricultural technologies is therefore a primary concern for policy makers. Unfortunately, the adoption of these productive technologies has proven to be sluggish across Kenya and Sub-Saharan Africa in general (World Bank, 2017). This leaves policy makers with a puzzle: on one hand, the lack of effective production is an existential threat, on the other, farmers simply seem unwilling to adopt new technologies that would not only improve their own food security, but would help to alleviate poverty more broadly.

Farmers may not adopt technologies for a variety of reasons. Economic theory suggests that non-adoption may be driven by the fact that farmers simply do not have sufficient information about different technologies and thus are unable to gauge their true productivity (Conley and Udry, 2010). As a result, policy makers have sought to implement interventions aimed at information provision with the the hope that increased information will diminish the uncertainty surrounding new technologies and induce take up.[1]

A body of literature, perhaps most famously illustrated with the example of preventing intestinal worms amongst schoolchildren in Kenya (Miguel and Kremer, 2004), has shown that the efficacy of these interventions is not determined only by their direct effect on the treated, but also the indirect effect on those who are not. Therefore, in evaluating policy, it is important that one considers treatment effects not only on those that are treated but also their neighbours and peers.

Social networks have been hypothesised as a candidate explanation for a variety of economic phenomena.[2] A social network is defined by individual members - nodes, and the social connections between them - links. Information about goods, services and ideas may flow between nodes through their links; a pair of farmers may discuss their experiences with a new banana cultivation technology, a village leader may inform his acquaintances about the benefits of modern stove technologies (Miller and Mobarak, 2014). If individuals are uncertain about the returns of different technologies, networks may provide an avenue through which they 'learn' about them. The policy maker has much to gain from understanding the role social networks play in diffusing technologies.

This idea has motivated a burgeoning body of literature that seeks to quantify social network

[1]Studies that involve technology training programs include: Foster and Rosenzweig (1995), Bandiera and Rasul (2006), Maertens and Barrett (2012), Islam et al. (2018), Chowdhury et al. (2019)

[2]For overview see Jackson (2010), Jackson et al. (2017)

effects of interventions in development settings. Most of these studies focus on identifying effective ways to target key individuals whom new information should be allocated to in order to maximise the diffusion information about new governmental policies (Alatas et al., 2019; Banerjee et al., 2018), products such as microfinance (Banerjee et al., 2013) and agricultural technologies (Beaman et al., 2018; BenYishay and Mobarak, 2018). These studies have not however, identified the effect of social networks when treatment is *randomly* allocated to network members. This gap in the literature has policy consequences; network data is expensive, difficult and may even be infeasible to collect in many development settings. Policy makers therefore, may be interested in implementing interventions where they must choose, for example, the fraction of a village which will be invited to an information session as opposed to specific 'important' farmers. Currently the evidence available to inform these decisions is limited.

We collected social network data from a set of 90 villages in Kirinyaga County, Kenya as part of a randomised controlled trial (RCT) seeking to investigate the following research question: what are the role of social connections between treated and untreated farmers in facilitating technology adoption amongst those that are untreated? Additionally, we seek to characterise the mechanisms through which these social network effects take form. In doing so we intend to broaden the suite of policy options available to decision makers when faced with the task of identifying the best methods to diffuse technologies within rural communities. This adds further depth to analyses which have previously identified the existence of spillover effects between treated and untreated farmers but has either not considered the role of social networks, or considered them in contexts where information is non-randomly allocated.

2 Literature Review

2.1 Previous Work

2.1.1 Learning From Family and Friends

A key barrier to technology adoption is uncertainty about the returns of adopting new technologies. Whilst individuals may directly learn about technologies from interventions such as information sessions (Chowdhury et al., 2019), they may also learn about them through their interactions with individuals they are socially connected with. Social networks provide researchers with a framework that allows them to utilise information about social connections in order to explain social and economic behaviour (Jackson, 2011). This idea was pioneered by Foster and Rosenzweig (1995) who found that neighbours' experience increased both the likelihood of adoption and the profitability of High Yielding Variety seeds in rural India.

Closely related is Conley and Udry (2010). This study provides a key insight into the potential mechanisms underlying social network effects. The authors find evidence that farmers adjust their production inputs to align with farmers that were in their 'information neighbourhoods' and that experienced above average crop yields in previous periods. This suggests that farmers learn about technologies not only through their own trial and error, but through observation of farmers around them.

In many development settings, including rural Kenya, home production is centralised and communities are organised along lines of familial connection (Iyer and Weeks, 2009). It is likely that farmers would learn about technologies from their network of family and friends. For example, Bandiera and Rasul (2006) finds evidence that the adoption decisions of friends and family are positivley associated with the adoption of sunflower cultivars amongst farmers in Northern Mozambique. In the absence of social network data, this study utilises geography as a proxy for social connections - an approach that was also utilised in earlier work such as Foster and Rosenzweig (1995) and the motivation for the analysis in Conley and Udry (2010).

The importance of geography in determining who one is socially connected with is well supported in the network formation literature (Crane-Droesch, 2018; Fafchamps and Gubert, 2007). Utilising geography as a proxy can be limiting in the sense that influential individuals within a network need not reside in a close proximity to others. Community leaders, for example, may not live in densely populated parts of a village or a slum. These limitations are summarised by Maertens and Barrett (2012) who finds that social connections between cotton farmers in India are better explained by features such as relative risk preference and common membership of caste. This has motivated the collection of richer sources of data - social network data. The collection of social network data typically involves asking respondents questions about the existence and/or the nature of their relationship with other respondents from the sample.[3]

2.1.2 Influential Agents

Social network data has motivated the idea that agents within a network are not uniformly influential – some may be more effective at spreading information than others. This has lead to the pursuit of answers to questions such as whether the decisions of particular agents disproportionately affect their network neighbour's decisions. Recently Alatas et al. (2019) has documented how 'tweets' of celebrities may contribute to the effectiveness of immunisation programs. Opinion lead-

[3]See Chandrasekhar and Lewis (2011) for various sampling techniques.

ers may also exist locally. Miller and Mobarak (2014) finds that village-level adoption of modern stove technologies in Bangladesh are correlated with the decisions of opinion leaders such as village elders and local politicians. The study finds that opinion leadership is most influential when the costs and benefits of technology are not well observed or understood. It also finds evidence that treated opinion leaders who end up not adopting technologies may create negative perceptions about technologies.[4]

2.1.3 Network Targeting

A natural question to ask is whether it is possible for policy makers to identify which agents are influential prior to an intervention. A seminal piece of work by Banerjee et al. (2013) finds that the 'diffusion centrality'[5] of first informed village leaders is positively associated with village-level adoption rates. The authors estimate a structural model of information diffusion to identify differences in 'endorsement' and 'information' effects[6] and find that the centrality of households in Indian villages first informed about microfinance, significantly helps to predict village-level take up.

Beaman et al. (2018) seeks to examine the conditions under which network theory[7] based targeting can outperform traditional methods of identifying influential individuals. In a similar flavour to Banerjee et al. (2013), the authors construct a structural model[8] of information diffusion which they then apply to data collected as part of a randomised field experiment focused on agricultural extension services in Malawi. The idea here is to determine the most effective 'injection points' according to the diffusion model, inform those individuals as part of an agricultural extension intervention, and then compare the adoption rates in villages that were subject to this method to ones where injection points were selected via other means. The study finds convincing evidence that identifying key individuals through running simulations of diffusion models outperforms other methods of targeting such as identifying village leaders or random allocation.

[4] We explore this idea in Section 5.1.3

[5] Diffusion centrality is a measure that combines features of degree and eigenvector centrality measures. See the supplementary material of Banerjee et al. (2013) for more details

[6] The authors distinguish between endorsement and information effects by noting that individuals may be influenced by their social connections in different ways. Firstly, social connections may be impact household adoption decisions in the sense that their connections, if informed, will give them information about microfinance with some probability. This is independent of whether or not the informants actually adopt the product themselves. Secondly, the endorsement effect stems from the fact that individuals who are connected with informants who actually take up the product themselves, are more likely to adopt microfinance than those who know informants that do not

[7] The models employed in Beaman et al. (2018) are 'simple contagion' and 'complex contagion' models.

[8] The model that the authors employ is known as a 'threshold' model. It is predicated on the observation that in some circumstances, individuals will adopt a new technology only if a sufficient number of their social connections are informed about it

The key question underlying BenYishay and Mobarak (2018) is how social network effects can be leveraged for more effective policy relating to technology adoption. This paper argues that opinion leaders who are properly incentivised will be more effective at encouraging technology adoption within a network. Echoing the findings found in Conley and Udry (2010) - that maize farmers in Malawi are appear most influenced by informants who they share a group identity, or face comparable conditions with. Once again, this study focuses on the incentives of 'lead farmers'; individuals who have been identified as potentially influential *a priori*.

2.1.4 The Random Provision of Information

Whilst studies such as Banerjee et al. (2013), Beaman et al. (2018) and BenYishay and Mobarak (2018) have found convincing evidence that targeting key network members can significantly improve the efficacy of interventions, these approaches necessitate the use of either detailed network data or local's knowledge about influential figures within communities. The focus on network targeting has yielded clear policy implications. If the adoption of productive technologies is to be encouraged across the developing world, policy makers should identify which individuals are influential or before an intervention, target them with information, and then rely on adoption practices to diffuse through their social networks. What if policy makers do no not have access to this information *a priori*? This is certainly not a trivial concern, and the difficulties with obtaining adequate network data have been well documented (Maertens and Barrett, 2012; Chandrasekhar, 2016; Manski, 1993). It may be too expensive or simply infeasible to collect social network data in many development settings. Randomly allocating treatment may be the only option policy makers have at their disposal.

Adding to this, there has been some work identifying conditions under which random allocation of information to network members may actually be more effective than targeting. Akbarpour et al. (2018) explores the notion that in some models of social diffusion,[9] randomly allocating information to a number of households may be more effective at spreading information than selecting to seed a lightly smaller amount. Empirically, Banerjee et al. (2018) finds evidence that there a circumstances under which the efficacy of an intervention would be improved if 'broadcasting' information randomly as opposed to targeting key individuals were utilised.

[9] For example, the popular Edros-Renyi model, the diffusion model from Banerjee et al. (2013) or the simple and complex contagion models discussed in Beaman et al. (2018).

2.1.5 Contributions

As it stands there is very little evidence that sheds light onto the role of social networks in diffusing technology when treatment is randomly allocated. A study by Islam et al. (2018) utilises an experimental design very similar to our own. Seeking to explain the adoption of modern rice farming practices in Bangladesh, it found evidence that the proportion of households invited to a training program were positively related to the adoption decisions of households that are not selected to participate in the training programs. Whilst the explanatory variable of interest, the exposure rate, captures the proportion of the sample which is invited to a training program, there is little consideration paid to the role the structure of each village social network specifically, plays in the adoption decisions of those households. Our work seeks to contribute to this gap in the literature by specifically investigating the role of social networks in facilitating the diffusion of technology when information about the technologies is provided randomly as opposed to targeted to specific network members. Our hope is to broaden the suite of options available to policy makers who wish to maximise these outcomes but do not have a evidence on alternatives to network targeting.

Aside from the policy implications of studying the social network effects associated with random allocation of information, there is an methodological advantage of this approach. Social network effects are fundamentally difficult to identify. The peer effects literature has illuminated a problem whereby spurious correlations between outcomes and network variables can be observed because of problems such as homophily,[10] and the existence of common shocks that affect all network members (Jackson, 2011). The collection of these identification problems has been dubbed the 'reflection problem' (Manski, 1993). In order to bypass these problems, papers such as Bandiera and Rasul (2006), Conley and Udry (2010) and Banerjee et al. (2013) rely on non experimental evidence for their examination of social network effects. The approach outlined in both this study and Cai et al. (2015) differs in the sense that the argument for the identification of network effects stems from random allocation of information. These approaches may act to buttress each other and provide further levels of external validity.

[10] Homophily as it relates to network theory is the notion that network members may choose to be connected because they share common (often, unobserved) preferences. This causes a problem for the identification of network effects in the sense that it may not be possible to determine whether network peers have been 'influenced' by each other, or the outcome of interest is being driven by some confounding factor (e.g. extraversion or resourceful) which also predicts their social connection.

3 Experimental Design

We utilise a 'cluster' RCT with randomisation levels at both the village and the farmer level. At the first stage of randomisation, 30 of the 90 sample villages were assigned to be pure control villages. Treatment is an invitation to an information session. These information sessions, carried out by the East Africa Market Development Associates (EAMDA),[11] involved invited farmers being liaised with an EAMDA instructor who provided information about the practical implementation and benefits of Tissue Culture Banana (TCB) technology which helps to eliminate the damage posed by pests and disease amongst banana crops.[12] Figure A4 in the Appendix displays the location of sample villages.

None of the farmers within pure control villages were treated.[13] Of the remaining villages, 15 were allocated to each of the treatment intensity groups in which either 20%, 40% 60% or 80% of farmers were invited to information sessions. Within each of these treatment villages a second round of randomisation took place to determine which of the sample farmers would be invited to the information sessions; the number of invitees within each village of course corresponded with which of the treatment intensity groups the village was assigned to in the first round of randomisation. Figure 1 illustrates the randomisation design.

We define a *treated* farmer as a farmer who resides in a treatment village and is invited to the information session. We refer to untreated farmers within treatment villages, that is farmers who don't receive training, as *spillover* farmers. To clarify, farmers in our dataset can be placed into 3 mutually exclusive categories, pure control, spillover and treated. The amount of treated and spillover farmers within each individual village is dictated by the village level randomisation which varies the the proportion of farmers that are invited to the information sessions.

1. *Control:* Farmers in villages that do not receive any treatment
2. *Treated:* Farmers within treated villages that receive treatment directly
3. *Spillover:* Farmers in treated villages that are untreated.

To clarify, if a village which contained 50 sample farmers was assigned to the 20% group, 10 farmers would have been invited to a information session; 20 if it were assigned to the 40% group and 0 if it were assigned to the pure control group.

[11]EAMDA is a consulting firm that offers business coaching and enterprise development in Eastern Africa.
[12]Further details about the intervention can be found in Chowdhury et al. (2019).
[13]Although these farmers were not treated, there is the possibility that there would be TCB adoption since TCB was not a 'new' technology.

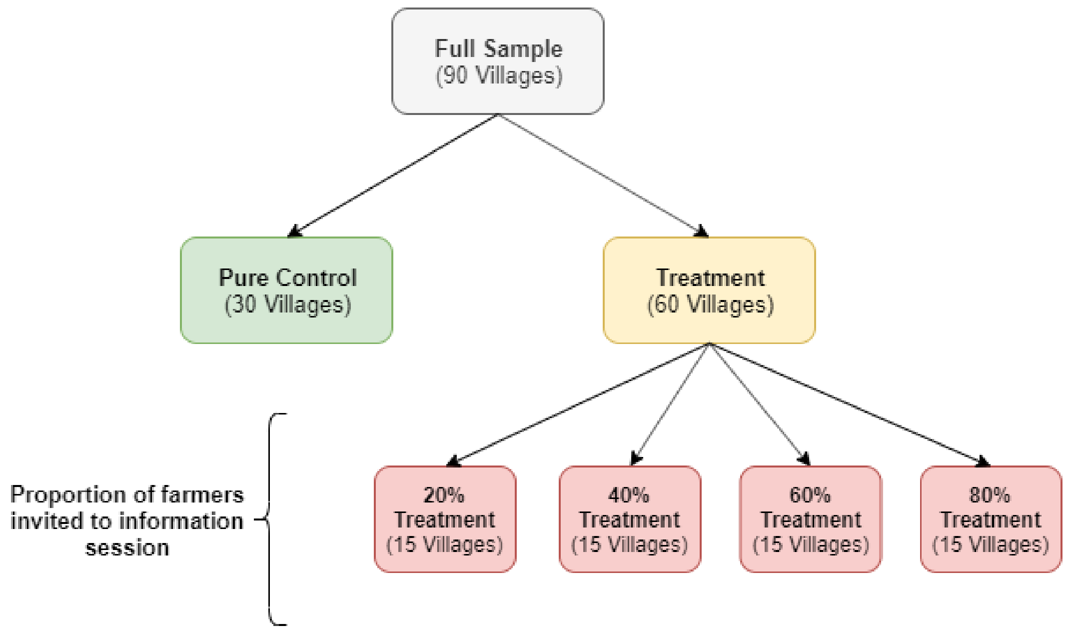

Figure 1: Randomisation Design

A key aspect of our experimental design is the fact that within treatment villages there exist both treated and untreated farmers. This allows us to identify not only the direct impact of the intervention on treated farmers, but also the effect of the intervention on those who are untreated. Specifically, we are able to assess the role that social connections between the treated and spillover farmers has on spillover adoption rates. A program evaluation based on this data is available in Chowdhury et al. (2019).

3.1 Defining a Social Network

3.1.1 Social Networks: Characterisation

Social networks are made up of agents (nodes) the social connections between them (links/edges). A network is observed when one observes these nodes and links. What constitutes a node depends on the context of a study and these have been defined as individuals, community groups (Ewert and Sunder, 2011) and even commercial entities (Fafchamps and Quinn, 2013). We define an agent as being a farmer within the sample. In terms of what constitutes a link, we adopt a definition whereby two farmers are connected if the respondent of one farmer responds affirmatively to a survey question asking whether they know a respondent from another farmer. This definition is consistent with much of the literature exploring the role of social networks on development outcomes (Conley and Udry, 2010; Banerjee et al., 2013; Miller and Mobarak, 2014; BenYishay

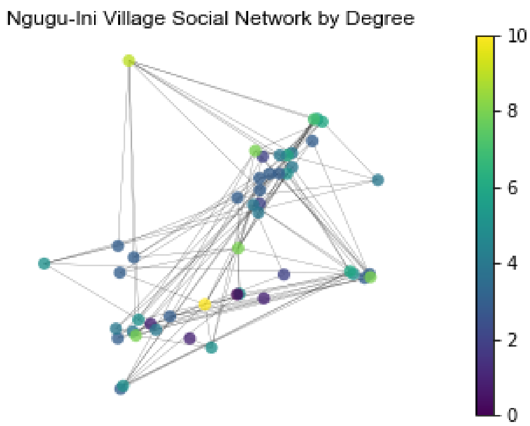

Figure 2: An example of a social network. Nodes are coloured by degree.

and Mobarak, 2018; Beaman et al., 2018). The definition we have introduced here implies that we are working with an *unweighted* and *undirected* network. We define a social network as existing within every village that we sample, with no crossover between villages (ie. respondents in one village may only be connected with farmers that were sampled from that village).

The economic relevance of social networks is well understood. Mbugua et al. (2019) describes social networks as mechanisms whereby information about goods, services and ideas flow. Jackson (2011) highlights the ubiquity of social networks pointing out that they serve as a candidate explanation for phenomena relating to labour markets, social learning processes and human capital decisions. More specifically, analysis of social networks may yield interesting insights into the drivers of technology adoption if individuals' adoption decisions are influenced by their social connections. Technical definitions of social networks provide a way to quantify the effect of these social connections.

3.1.2 Social Networks: Technical Defintion

A network is a pair $G = (V, E)$[14] where V is a set of *nodes* and E a set of *edges* representing connections between them. A link between two nodes is defined when there exists an affirmative response from one agent to knowing another. We do not consider the 'direction' of this link in the sense that agent i being linked with j implies that agent j is linked with agent i. This is an example of an *unweighted* and *undirected* network. We characterise these connections with an adjacency matrix $A := A(G)$. Since the network we consider is unweighted and undirected, this adjacency matrix is made up of 0s and 1s where $A_{ij} = \mathbf{1}\{ij \in E\}$, meaning that $A_{ij} = 1$ if individuals i and

[14]We borrow our notation from Chandrasekhar and Lewis (2011)

j are linked and 0 otherwise. For the purposes of our analysis, we are interested in characteristics of social networks (see Section 3.5). These are functions $w(G)$, of the network G. In practice, it is useful to compute these characteristics by considering the adjacency matrix $A(G)$. A visualisation of a social network is given in Figure 2, it represents the social network for a particular village in our sample: Ngugu-Ini.

3.2 Data Collection

Along with the information sessions, we conducted a survey which is described in Chowdhury et al. (2019). In this survey, treated, untreated and farmers in the control villages were asked questions about, for example, their income, but also about their social networks. These were carried out in three rounds at baseline (May-June 2016), midline (Oct-Nov 2017) and endline (Oct-Nov 2018) time periods. The baseline survey was used to collect farmer characteristics such as sex, age and education of the respondent farmer, household size, number of plots owned by the farmer, number of rooms in the house, access to electricity, ownership of livestock, access to cash saving, and access to radio and television.

In total, 4719 farmers were sampled across the 90 target villages. On average, 53 farmers in each village were sampled, this figure is slightly higher than the 50 farmer target. Finally, at endline (Oct-Nov 2018) a follow up survey conducted aimed to ask all sampled farmers the same social network related questions as in midline. A total of 4101 farmers responded to this survey. The only difference between this survey and midline was that social network questions were only asked about individuals which the respondent had indicated being linked with in the midline survey. For example, if farmer a reported knowing an farmer b but not farmer c, farmer a would only be asked social network questions regarding farmer b at endline.

Overall baseline, midline and endline information is available for about 87% of farmers who were initially surveyed. Summary statistics provided in Table 2 indicate that baseline characteristics are balanced across groups. Attrition is also unrelated to any of these, evidence for which can be found in Chowdhury et al. (2019). We also test for the association between network degree and attrition and find that they are unrelated (Table A10 in the Appendix).

Table A9 in the Appendix presents the number of observations we obtained for each group with and without controls. The specific controls we collected are discussed in Section 3.4.

Table 1: Adoption Rates by Treatment Status

	(1) Midline	(2) Endline	(3) Δ
Control	10.80 %	27.06 %	16.26%
Spillover	12.53 %	27.61 %	15.08%
Treated	14.96 %	37.94 %	22.98%
$N = 4101$			

3.2.1 Social Network Questions

At midline and endline, participants were asked a set of social network questions such as *"Do you know [...]?", "Is [...] in any farmer group with you?"* and *"Have you discussed banana cultivation practices with [...]?"*. If there was no attrition, this would allow us to determine the existence of social connections between all pairs of farmers in sample villages.[15] We provide a sample of the questionnaire used to obtain social network data in Figure A3 in the Appendix.

3.3 Outcome Variables

Since we are interested in understanding the role of technology adoption, our key outcome variable is a binary random variable indicating whether or not a farmer has adopted TCB technology. We obtained observations of adoption decisions at two time periods: midline and endline. Hence, our outcome variables can be represented by the random variable $Adopted_{itv}$ which equals 1 if individual i in village v adopted TCB by period t where t refers to either midline or endline.

Interestingly, the change in adoption rates between midline and endline are comparable for spillover and control farmers. This is illustrated in Table 1. The consequences of this with relation to our analysis are explored in Section 5.

[15]The sampling method we employ shares similarities to the 'induced subgraph' method as discussed in Chandrasekhar and Lewis (2011) Our sampling method differs slightly in that we do not ask ask each of the n sample nodes about $n-1$ others. This is clarified in Section A.2.1, although the adjacency matrix generated by our sampling method would be the same as the induced subgraph method assuming there were no disagreeable links (i.e. situations where i reports knowing j but j reports not knowing i).

3.4 Controls

We have a set of baseline characteristics $\mathbf{x}_{iv}$ which contains farmer level information relating to socioeconomic factors such as livestock ownership and access to particular technologies[16] as well as information about past banana cultivation behaviours. This allows us to control for farmer specific characteristics which may contribute to adoption of TCB technology. Summary statistics for the controls are provided in Table 2.

Table 2: Village Level Controls By Treatment Status

	Control		Treatment		Diff.	
	Mean	Std. Dev.	Mean	Std. Dev.	Estimate	t-stat
Household Size	402.70	76.76	408.20	50.62	-5.50	-0.36
Land Size	198.57	61.73	208.67	74.49	-10.10	-0.68
Electricity Access	93.53	22.76	86.13	20.47	7.40	1.50
Cattle	102.03	18.01	96.78	12.38	5.25	1.44
Savings	1.60	2.01	1.23	1.82	0.37	0.84
ln(Household Income)	16.81	0.27	16.78	0.23	0.03	0.60
Observations	30		60		90	

As we can see from the right hand side columns in Table 2, there is no statistically significant differences in average values of control variables across pure control and treatment villages. Further balance checks and evidence that attrition is not related to any of these variables is available in Chowdhury et al. (2019). We note that there are 3998 farmers for which we observe a full set of individual-level controls. Hence, whilst we sample 4101 total farmers any estimation results on the full sample with controls contains 3998 observations. We chose these particular controls simply because of the fact that these were the ones collected as part of the survey.

3.5 Network Variables

The 'effect' of a social network is hard to grasp in totality; characteristics are used to pin down features of social networks that may be of interest to researchers. Farmers with a relatively high network degree for example, are more 'well connected' since they have more social connections.

[16]For full list see Table 2.

3.5.1 Degree

The network degree of the farmer i is simply a count of the links it shares with the other farmers in the sample. We utilise two of these counts, one referring to the amount of links with farmers in the full sample and another referring to the amount of links with treated farmers.

1. *Degree:* Number of farmers the farmer i is connected with.

$$Degree_{iv} = \#Links\ within\ village_{iv}$$

2. *Degree of Treated Farmers:* Number of treated farmers the farmer i is connected with. This measure is only observable for spillover sample farmers.

$$Degree^{T}_{iv} = \#Links\ with\ Treated\ Farmers_{iv}$$

It is important to note that the first of these is endogenous to our outcome variables. This is because of a concept known as homophily. Homophily is the idea that network nodes may be connected because of some unobserved preference structure, and thus any correlation between network measures and outcome variables may be spuriously driven by that unobserved preference. An example used by Manski (1993) refers to this idea in the context of identifying the peer effects of smokers. It is not possible to identify whether smokers smoke because they are influenced by their friends, or because one is more likely to be friends with those with whom they share common personality traits which are conducive for smoking (e.g. tendency to engage in risky behaviour). This problem motivates the definition of the second measure, which is less likely to suffer from the endogeneity problem since the variation is driven by treatment assignment which of course, is random. Evidence for this correlation is provided in Table A16 in the Appendix.

For robustness we also conduct analysis utilising a slightly different measure known as 'density'. We report results in the appendix. The only difference between degree and density is that density measures are weighted to reflect the size of the network that is considered. Formally: *Link Density:* Proportion of farmers the farmer i is connected with. *Link Density of Treated HHs:* Proportion of treated farmers the farmer i is connected with. These measures are prevalent in studies such as Cai et al. (2015). Our results, discussed in Section 4 using degree as an explanatory variable corroborate those using density as an explanatory variable which are displayed in the Appendix.

In our exploration of potential mechanisms (Section 5) we introduce a number of other explanatory variables. These are counts of links with particular sub populations (e.g. number of links with

Treated Non Adopters) and are very similar to the definitions provided above.

4 Estimation Results

4.1 Main Results

In this section we are interested in answering our key research question: what is the role of social connections between treated and untreated farmers in facilitating technology adoption amongst those that are untreated?

We provide a set of preliminary results in the Appendix (see Section A.3). Firstly, they provide evidence that treatment has a direct effect on farmers that are treated. Secondly, they provide some evidence of the association between higher levels of degree amongst farmers and overall village-level adoption of TCB.

4.1.1 The Effect of Degree

In order to examine how the number of one's social connections affects their adoption decisions, we test for an association between degree and the adoption of TCB. We estimate Equation (1) on the spillover, treated and control groups.[17]

$$
\begin{aligned}
Adoption^G_{itv} = \beta_0 + \beta_1 Degree_{iv} + \beta_2 Treatment\ Level_v \\
+ \beta_3[Degree_{iv} \times Treatment\ Level_v] + \gamma' \mathbf{x}_{iv} + u_{iv}
\end{aligned}
\tag{1}
$$

Where $Adoption^G_{itv}$ equals 1 if individual i, who is a member of group G, in village v, adopts TCB by period t. The vector $\mathbf{x}_{iv}$ is a set of controls. Our parameters of interest are (β_1, β_3) which capture the effect of having more social connections and interactions with treatment levels respectively. We report estimation results in Table 3.

Columns 1 and 2 of Table 3 suggest that control farmers are more likely to adopt TCB if they are connected with an additional farmer in midline. In contrast, we observe no effect amongst treated farmers. For farmers within the spillover sample, Columns 7 and 8 indicate that an additional social connection increases the probability of TCB take up by 1.6 percentage points. The results in Columns 9 and 10 indicate that this effect is invariant to the intensity of treatment within

[17]The superscript G is used to denote adoption outcomes for each of these groups: $G \in \{S, T, C\}$

villages. The increased likelihood of adoption associated with being connected with an additional farmer is no bigger for those in villages where 20% of farmers are treated, than in villages where 80% are treated. Although insignificant in the control group, we find evidence that the effect size of knowing an additional social connection is comparable for control and spillover farmers.

Table 3: The Effect of Degree on Adoption Outcomes by Group

	Control		Treated				Spillover			
	(1) (M)	(2) (E)	(3) (M)	(4) (E)	(5) (M)	(6) (E)	(7) (M)	(8) (E)	(9) (M)	(10) (E)
Degree	0.00946**	0.0111	0.00390	0.00701	-0.00745	-0.00162	0.00692**	0.0155***	0.00700	0.0166***
	(0.00419)	(0.00659)	(0.00363)	(0.00632)	(0.0118)	(0.0222)	(0.00343)	(0.00481)	(0.00544)	(0.00541)
Int.=40 × Degree					0.0171	0.00286			-0.000386	0.000360
					(0.0138)	(0.0256)			(0.00808)	(0.0128)
Int.=60 × Degree					0.00445	0.0211			-0.000491	0.00226
					(0.0137)	(0.0247)			(0.00979)	(0.0127)
Int.=80 × Degree					0.0175	0.00474			0.00114	-0.0146
					(0.0124)	(0.0239)			(0.0104)	(0.0127)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1337	1337	1280	1280	1280	1280	1381	1381	1381	1381

Notes: Constant and treatment levels omitted from output. (M): Midline, (E): Endline.
Standard errors clustered at the village level
$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

These results are unsurprising in the sense that we would expect farmers who receive treatment directly to be less influenced by their social connections. Assuming that signals from trained EAMDA instructors are less noisy than those of other farmers, this result accords with much of the theory surrounding social learning.[18] Conversely, amongst farmers who do not receive any information provision and thus may be able to receive signals from their network neighbours, we find a significant, positive effect of knowing an additional farmer. What *is* surprising however, are the comparable effect sizes between spillover and control farmers. We explore this result further in Section 5.

Whilst we provide evidence that a greater number of social connections is associated with a higher propensity to adopt TCB by endline, these results merely indicate a correlation between social connections and adoption outcomes since the network variable is endogenous.[19] Moreover, these

[18] The intuition these models is that farmers first begin with a belief the underlying distribution of returns for a particular technology. They update their beliefs ('learn') by conditional on noisy signals produced by network neighbours. The point here is that treated farmers have no need to update their beliefs given that they received direct training surrounding the use of TCB. For a concrete example see for instance Conley and Udry (2010)).

[19] Network variables are endogenous in the sense that there may be some unobserved factors which make a farmer both more likely to have more social connections but also likely to adopt the technology. farmers who have similar propensities to adopt new technologies may be more likely to be connected, this may be driven by some underlying preference structure which we do not observe. This phenomena is known as 'homophily' in the network literature. We provide an example of this in Section 3.5.

findings do not shed light on the significance of the social connections between treated and untreated farmers specifically.

4.1.2 Popularity and Diffusion Effects

Whilst establishing that individuals who have more social connections are more likely to adopt TCB is an important starting point for our analysis, the finding does not shed light onto the significance of social connections between treated and untreated farmers specifically. Understanding this relationship is meaningful from a policy perspective, as it allows one to more clearly characterise the effect of an additional unit of treatment by taking into account it's diffusion through the social network.

We suggest that the social network effects we observe may be classified into two broad categories:

(i) *The "popularity" effect*: the likelihood of adoption being associated purely with the number of farmers one is connected with. This is irrespective of whether or not those farmers have been treated.

(ii) *The "diffusion" effect*: the likelihood of adoption being associated with the number of *treated* farmers one is connected with.

This distinction is important. If the effect of the social network manifests as a popularity effect but not a diffusion effect, it would imply that the association between social connections and TCB take up is independent of treatment. If this were the case, it could be that social networks do not necessarily facilitate the diffusion of treatment. In the following section we test for the existence of the diffusion effect.

4.1.3 The Diffusion Effect: Connections Between Treated and Spillover Farmers

To test for the existence of a diffusion effect, we examine the relationship between the number of treated farmers one is connected with and the adoption of TCB. We run a specification which is analogous to Equation (1) however, that utilises a count of the number of treated farmers, farmer i is connected with. We estimate the following specification on the set of spillover farmers.

$$\begin{aligned} Adoption^S_{itv} = \beta_0 + \beta_1 Degree^T_{iv} + \beta_2 Treatment\ Level_v \\ + \beta_3[Degree^T_{iv} \times Treatment\ Level_v] + \gamma' \mathbf{x}_{iv} + u_{iv} \end{aligned} \tag{2}$$

Where $Degree^T_{iv}$ refers to the number of treated farmers spillover farmer i is connected with, $\mathbf{x}_{iv}$ is a set of controls and our parameters of interest are (β_1, β_3). Unlike the specification in Equation (1), these results can be interpreted as causal estimates. Treatment is randomly allocated, hence $Degree^T_{iv}$ is exogenous in the sense that it varies with the treatment level of v. Evidence for the correlation between treatment level and $Degree^T_{iv}$ is reported in Table A16 in the Appendix.

Our results for this specification are outlined in Columns 1-4 of Table 4.[20] Column 2 contains one of our key findings, that knowing an additional treated farmer increases the likelihood that an untreated (spillover) farmer adopts TCB by 2.4 percentage points.

Although we have argued that the network variable $Degree^T_{iv}$ is exogenous, we still face an identification issue. Recall, we defined the popularity effect as the association between the likelihood a farmer adopts TCB and their total number of social connections. We found evidence for this in Section 4.1.1. If the popularity effect exists, then there is a high likelihood that we would observe a diffusion effect since by virtue of randomisation, farmers that have many social connections will be connected with many treated farmers. We cannot directly disentangle these effects, and it is possible that treated links do not diffuse any more information about TCB than untreated ones. The positive association between treated links and adoption may be an artefact of the fact that farmers are simply more likely to adopt TCB if they have more friends, irrespective of their friends' treatment status. This is problematic from a policy perspective as finding evidence for a non-existent diffusion effect would overplay the efficacy of this particular intervention.

4.1.4 A Counterfactual: Pseudo Treatment

To address the identification issue put forward in Section 4.1.3 we construct a counterfactual using the control group. The counterfactual is predicated on the following idea. Suppose that the diffusion effect exists. We would expect links between treated and untreated farmers to be predictive of adoption for those that were untreated. We find evidence for this in Section 4.1.3. Now, if this effect were not being driven by the popularity effect, then links between farmers randomly allocated into 'pseudo-treated' and 'pseudo-spillover' groups in control villages should not be predictive of adoption amongst farmers, since treatment is absent in these villages. If these links were associated with a positive increase in adoption outcomes, it would imply that adoption outcomes depend on social connections irrespective of their treatment status and confirm that the popularity effect is driving our results.

In order to test this, we randomly allocate pure control farmers into 'pseudo-treated' and 'pseudo-

[20]The results in Column 7 and 8 are discussed in Section 4.2

spillover' samples. This allocation replicates the process used in actual treatment villages, the only difference being of course, that these farmers did not actually receive treatment. A more thorough exposition of the method we use is available in Section A.5 in the Appendix. Once allocated, we are able to obtain the number of pseudo-treated farmers each pseudo-spillover farmer is connected with. We then estimate Equation 3 which is analogous to Equation 2[21]. The resulting estimates are reported in Table 4 Columns 5 and 6.

$$\begin{aligned} Adoption^C_{iv} = \beta_0 + \beta_1 Pseudo\ Degree^T_{iv} + \beta_2 Pseudo\ Treatment\ Level_v \\ + \beta_3 [Pseudo\ Degree^T_{iv} \times Pseudo\ Treatment\ Level_v] + \gamma' \mathbf{x}_{iv} + u_{iv} \end{aligned} \tag{3}$$

Where $Adoption^C_{iv}$ refers to the adoption decisions of the pseudo-spillover sample. $Pseudo\ Degree^T_{iv}$ refers to the number of pseudo-treated farmers the farmer i is connected with and $\mathbf{x}_{iv}$ is a set of controls. Our parameters of interest are (β_1, β_3).

Table 4: Number of Links with Treated Farmers on Adoption Outcomes of the Spillover Sample

	Spillover				Control (Pseudo-Treatment)			
	(1) (M)	(2) (E)	(3) (M)	(4) (E)	(5) (M)	(6) (E)	(7) (M)	(8) (E)
Links (Treat. HHs)	0.00487 (0.00533)	0.0240*** (0.00728)	0.00257 (0.00921)	0.00951 (0.0114)				
Int.=40 × Links (Treat. HHs)			0.0203 (0.0141)	0.0402*** (0.0139)				
Int.=60 × Links (Treat. HHs)			-0.00616 (0.0124)	0.0243 (0.0197)				
Int.=80 × Links (Treat. HHs)			-0.0149 (0.0115)	-0.00753 (0.0238)				
Links (Pseudo Treat. HHs)					0.000551 (0.00570)	0.0140 (0.0113)	0.00655 (0.00965)	-0.00277 (0.00934)
Pseudo Int.=40 × Links (Pseudo Treat. HHs)							-0.00431 (0.0119)	0.0341 (0.0241)
Pseudo Int.=60 × Links (Pseudo Treat. HHs)							-0.00974 (0.0178)	0.0230 (0.0242)
Pseudo Int.=80 × Links (Pseudo Treat. HHs)							-0.0258 (0.0165)	0.00524 (0.0280)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1381	1381	1381	1381	671	671	671	671

Notes: Constant omitted from output. Standard errors clustered at village level. M: Midline, E: Endline. Constant omitted from output. $^* p < 0.10, ^{**} p < 0.05, ^{***} p < 0.01$

We find no evidence that links between pseudo-treated and pseudo-spillover houses are predictive of

[21]Figure A1 in the Appendix confirms that the distributions of actual and pseudo link densities are comparable.

adoption amongst pseudo-spillover households. We have evidence that in the absence of treatment, links between farmers randomly allocated into different groups do not account for any increase in the likelihood of take up. It is therefore likely, that the diffusion effect we observe in Section 4.1.3 is not simply an artefact of the popularity effect, whereby farmers are more likely to adopt simply because they are connected with more network members, irrespective of their connection's level of treatment.

We have shown that connections with treated farmers have a significant affect on adoption outcomes even when these treated farmers are not targeted or identified as influential. This has important implications for policy. The result not only solidifies our key finding - that connections between treated and untreated households are predictive of adoption amongst those that are untreated, it also suggests that social networks may help to proliferate technologies even when information is randomly allocated.

4.2 Interactions with Treatment Intensity

Our experimental design allows us to identify whether social network effects interact with the intensity of treatment. As we have intimated through this section, there is mixed evidence about whether the intensity of treatment determines the magnitude of social network effects. We will discuss this explicitly here.

With respect to degree, it appears as though there is no evidence of there being a difference across treatment levels in the association between the number of links and the likelihood of adopting TCB (See Table 3, Columns 9 and 10). One explanation for this may be the fact that the distribution of degree is heavily right skewed (See Figures A1 and A2 in the Appendix). If farmers, in general, do not have many social connections, it is unlikely that increasing information provision is going to increase the likelihood of take up since it is not necessarily likely that these farmers would be connected with more informed farmers.

On the contrary, our results where we test the association between the number of links with treated farmers and the adoption outcomes of the spillover sample, seem to suggest that social network effects may not be invariant to the level of treatment (refer to Equation 2 and Table 4). Column 4 of Table 4 indicates that the diffusion effect we described in Section 4.1.3 is greatest in villages where 40% of farmers are invited to information sessions. The reasons for this are beyond the scope of this paper however, this may be an area of interest for future research.

5 Possible Mechanisms and Discussion

5.1 Hypothesising Potential Mechanisms

Admittedly, the preceding results carry a slight contradiction. A key feature of our findings is the fact that adoption rates in the spillover and control groups are comparable (Table 1). Despite this, in Section 4.1.3 we find that there exists a diffusion effect: the causal association between the number of treated links a spillover farmer has and the adoption of TCB. We find that this is unlikely to be a spurious correlation since we find evidence that links between farmers randomly assigned to pseudo-spillover and pseudo-treated groups do not predict adoption in the control group.

One explanation is that farmers within the control group are are in fact connected with other farmers who, although untreated, still carry information about the technology. TCB is not a new technology, it is certainly possible that farmers connected with those who adopted before the midline survey experience in essence the same 'diffusion' effect albeit from previous adopters rather than treated individuals. Our hypothesis is that if farmers within the control group have relatively high levels of TCB adoption in baseline[22] (or indeed in midline for some reason independent of treatment), social connections to these adopter farmers may be as effective at diffusing technology than those in the treatment group.

This hypothesis would explain the high levels of adoption observed amongst control farmers and would also serve as a candidate mechanism underlying the diffusion effect discussed earlier. We also explore other mechanisms such as the possibility of social networks facilitating negative transmission channels and the importance of farming neighbours for the diffusion of technology, in order to provide some alternative explanations for the phenomena we observe.

In this section we investigate the following:

(a) Is there evidence of 'diffusion' without treatment?

(b) The importance of links with adopters

(c) The possibility of negative transmission channels

(d) The importance of neighbours and observability

[22]Baseline adoption could be though of as 'pre-treatment' adoption. Unfortunately, due to budget constraints, we do not have data for baseline levels of adoption.

5.1.1 'Diffusion' in the Absence of Treatment

As suggested, connections with adopters may create similar effects amongst those in control villages, as connections with treated farmers do amongst those in the spillover sample. We can think of these adopters as being a 'proxy' for treated farmers in control villages. If there exists a positive relationship between links with these adopters and take up of TCB, it may explain the association between degree and adoption we observe amongst control farmers. In order to examine whether this phenomena is present within the control group, we test for an association between links with midline adopters and endline adoption of TCB. We estimate the following on the control group:

$$Adoption\ in\ Endline_{iv} = \beta_0 + \beta_1\{\#Links\ w/\ Midline\ Adopters\} + \gamma'\mathbf{x}_{iv} + u_{iv} \tag{4}$$

The parameter of interest here, β_1, captures the association between connections with adopters and future adoption of TCB.[23] This of course, is purely an association and does not capture the 'effect' of these links per se, since there may be social traits that make adopters more likely to be linked with other adopters.[24] Results are reported in Table 5.

Table 5: Control Group: Links with Midline Adopters

	(1)	(2)
# Links with Midline Adopters	0.319***	0.307***
	(0.0323)	(0.0297)
Controls	No	Yes
Observations	1371	1337

Notes: Constant omitted from output. Standard errors clustered at village level. Outcome variable is whether one adopted in endline but not midline. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

We find strong evidence of a positive correlation between connections with midline adopters and the likelihood of TCB adoption in endline. This finding has two implications. Firstly, it suggests that if there were high levels of baseline or midline adoption amongst control farmers, the high levels of take up we see by endline may have been driven by non-adopters being connected with adopters. Secondly, it provides a candidate explanation for the existence of the social network effect in the control group – social connections are significant predictors of TCB adoption in that being connected with adopters increases the likelihood that a farmer adopts TCB themselves.

[23]The variable *Adoption in Endline*$_{iv}$, is different from the outcome variable we have discussed elsewhere. It equals one if the farmer did not adopt in midline, but then adopted in endline. This is contrast to *Adopted Endline*$_{iv}$ which equals one if the farmer adopted in *either* midline or endline

[24]This is the same reasons discussed in Section 4.1.1

5.1.2 Links with Adopters

The control group analysis in Section 5.1.1 sheds light onto a key potential mechanism – links with adopters. The diffusion effect may be driven by the fact that treated farmers are more likely to adopt TCB and that being connected with treated farmers makes one more likely to be connected with adopters. In order to investigate this, we test for the association between links with adopters and TCB take up across the full sample of control, treated and spillover farmers. We then formally test for differences in effect size between these groups by estimating a model with interaction terms.

We estimate:

$$\begin{aligned} Adoption_{iv} = \beta_0 + \beta_1\{\#Links\ w/\ Midline\ Adopters\} \\ + \beta_2\big[\mathbf{1}\{Spillover\ HH\}_{iv} \times \{\#Links\ w/\ Midline\ Adopters\}\big] \\ + \beta_3\big[\mathbf{1}\{Treated\ HH\}_{iv} \times \{\#Links\ w/\ Midline\ Adopters\}\big] + \gamma'\mathbf{x}_{iv} + u_{iv} \end{aligned} \tag{5}$$

Where $Adoption_{itv}$ equals 1 if farmer i in village v adopted TCB by time t. $\mathbf{1}\{Spillover\ HH\}_{iv}$ and $\mathbf{1}\{Treated\ HH\}_{iv}$ are indicators equaling 1 if the farmer belongs to either of the respective groups. The baseline group are the pure control farmers, and thus β_1 identifies the association between links with midline adopters and the adoption decisions of these farmers. We are principally interested in the set of parameters (β_1, β_2).[25]

We find strong evidence of an association between links with midline adopters and the likelihood that control farmers adopt TCB. Interestingly, since our estimate of β_2 is insignificant, we conclude that the size of this association is the same in both the control and spillover groups. We interpret this result as suggesting that the effect of knowing an additional midline adopter is the same for both control and spillover farmers, the magnitude of this effect is about 2.6 percentage points.

This result provides us with an important insight. Knowing an adopter is just as predictive for adoption in the control group as in the spillover group. The effect is independent of whether or not the farmer resides in a village where there is no treatment at all.

This suggests that social networks may facilitate the diffusion of TCB primarily through links with

[25] β_3 identifies the association between links with midline adopters and the outcome variables of treated farmers. It is interesting that we observe a statistically significant, positive result on this coefficient especially given that we do not observe any evidence of links being predictive of adoption in Table 3. This relationship is beyond the scope of this paper, however we note that it may be a guide for future research.

Table 6: Links with Adopters

	Pooled		Spillover	
	(1) (M)	(2) (E)	(3) (M)	(4) (E)
Links with Adopters	0.0145*** (0.00532)	0.0260*** (0.00740)	0.0158*** (0.00514)	0.0267*** (0.00693)
Spillover =1 × Links with Adopters	-0.000202 (0.00687)	-0.00183 (0.00928)		
Treated =1 × Links with Adopters	0.000477 (0.00641)	0.0205** (0.00919)		
Links with Treated Farmers			-0.00522 (0.00563)	0.00870 (0.00674)
Controls	Yes	Yes	Yes	Yes
Observations	3998	3998	1381	1381

Notes: Columns 1 and 2 estimate equation (5) on all farmers. Columns 3 and 4 restrict the sample to spillover farmers and estimates the effect of links with adopters holding the amount of treated links constant.
Constant omitted from output. Standard errors clustered at village level.
$^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

those who adopt the technology. The results we present in Table 4: that spillover farmers are more likely to adopt TCB by endline if they know more treated farmers, may in fact be driven by the fact that treatment has a direct effect on those who are treated. Treated farmers are more likely to be adopters (evidenced by Table A12) and spillover connections with these farmers are predictive of their own eventual adoption. This idea is consistent with our findings that there are high levels of endline adoption in the control group, and also explains why links with pseudo-treated farmers are not predictive of adoption: these farmers are no more likely to adopt than any other since there is no direct treatment effect.

5.1.3 Transmission of Negative Information

Whilst we have provided evidence that links with adopters are a key mechanism for the diffusion of TCB, we must also consider the existence of other candidate mechanisms. Social networks need not strictly transmit 'positive' information about a technology. In the same way that a farmer may transmit positive information about a technology in the event of a greater than average crop yield, they may also transmit negative information about it if it produces adverse outcomes.

One could posit that the relatively high levels of adoption in the control group are high not because of the diffusion effect, but because there are relatively low levels of adoption in the spillover group. This may occur for example, because negative information about the technologies were passed through social networks in the spillover group and not the control group. A plausible reason why this may have occurred is the fact that spillover farmers were exposed to some farmers while attending information sessions, decided not to adopt TCB and subsequently imparted bad impressions about the technology on their social connections.[26] This is less likely to have occurred in the control villages since there was no treatment, and presumably those adopting TCB would have positive views of it.

A key idea that is prevalent in the seminal paper by Banerjee et al. (2013) is a distinction between information and endorsement effects. The authors argue that whether a person participates in a microfinance program depends on a) whether the person is aware of the opportunity and b) whether the persons social connections participate in the program.[27] The former is known as the information effect and the latter is known as the endorsement effect.[28] Implicit in this distinction is the idea that there is heterogeneity in the signals one receives from their social connections; signals from those who are treated but do not adopt are far less correlated with adoption than signals from those who do. Also related is Miller and Mobarak (2014). This paper finds evidence that the non-adoption of modern stoves by village leaders created negative impacts on the likelihood that villagers in Bangladesh would take up the technology.

Our exploration of mechanisms relates to this literature by examining whether there is evidence that social networks facilitated the transmission of negative information about TCB. This analysis may help to explain the high levels of adoption apparent in the control group, but also clarify the importance of links with other adopters. Assuming that individuals who are treated but do not adopt a new technology are more likely to transmit negative rather than positive information to their peers, we are able to test the existence of negative information flows by investigating the relationship between the adoption outcomes of the spillover group and those who were treated by did not adopt. In order to test this, we estimate the following:

$$\begin{aligned} Adopted^S_{itv} = \beta_0 + \beta_1\{\#Links\ w/\ Treated\ Non\ Adopters\} \\ + Degree_{iv} + \gamma'\mathbf{x}_{iv} + u_{iv} \end{aligned} \tag{6}$$

[26]Furthermore, those invited to information session are more likely to spread bad information about TCB because in some sense they did not 'choose' to be exposed to it. Those that may have, for example, researched TCB independently would, we believe, be less inclined to transmit negative information since they may have simply chose not to adopt and did not incur the time cost of attending training.

[27]The authors note that this endorsement effect is actually a "catch all" for any interaction that provides more than simple information diffusion

[28]This draws some similarities to the distinction we put forward in Section 4.1.2. The difference being that our distinction does not depend on whether social connections adopt or not, rather we focus on treatment status.

Where $Adopted^S_{itv}$ is an indicator variable describing the adoption decisions of the spillover sample. The sub population "*Treated Non Adopters*" refer to those farmers who were invited to treatment but did not adopt the technology. $Degree_{iv}$ denotes the total number of social connection the farmer has. The parameter of interest β_1, captures the association between a spillover farmer's links with treated non adopters and their adoption of TCB conditional on the total number of social connections they have. Once again, this does not identify any causal effect and is simply an association. We report our results in Table 7.

Table 7: Spillover Links With TNAs

	(1) (M)	(2) (E)
Links with TNAs	0.00448 (0.00520)	0.000535 (0.00581)
Degree	0.00551 (0.00347)	0.0168*** (0.00473)
Controls	Yes	Yes
Observations	1304	1304

Notes: Constant omitted from output. Standard errors clustered at village level. TNA: Treated Non Adopter. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$.

If negative information about TCB were being passed along through social networks, we would expect that being connected with more farmers that were treated but did not adopt the technology would be negatively associated with adoption outcomes. The results in Table 7 seem to refute this notion. We find that, conditional on the number of total social connections, there is no effect on adoption from knowing an additional Treated Non Adopter.

These results do not necessarily suggest that negative information does not diffuse through social networks, rather they suggest that these at least do not appear to be coming from treated non adopters. The positive association that we observe may be driven by the fact that farmers who are connected with more Treated Non Adopters are likely connected with more adopters as well. Even if social connections with Treated Non Adopters were producing some negative effect with respect to adoption, it is clear that at least on average these are cancelled out.

5.1.4 Neighbouring Plots

In the spirit of Conley and Udry (2010),[29] we investigate whether the adoption decisions of neighbouring farmers impact the likelihood of ones own adoption of TCB. We define one's set of neighbours as those who share a neighbouring farming plot. Note that these may not necessarily be the neighbour that we conceive of conventionally,[30] and instead refer to 'farming neighbours'.

In order to test for an association between one's adoption of TCB and their neighbours, we estimate:

$$Adopted_{itv}^{G} = \beta_0 + \beta_1\{\#Neighbours\} + \gamma'\mathbf{x}_{iv} + u_{iv} \tag{7}$$

Where $Adopted_{itv}^{G}$ is an indicator for the adoption of TCB for individual i in village v and who is a member of group G adopts TCB by period t. The parameter of interest is β_1 which captures the association between an additional farming neighbour and TCB adoption. Results from our estimation are reported in Table 8.

Table 8: Number of Farmers With Neighbouring Plots

	Control				Spillover			
	(1) (M)	(2) (E)	(3) (M)	(4) (E)	(5) (M)	(6) (E)	(7) (M)	(8) (E)
All Neighbours								
# Neighbours	0.00440	0.00733			0.00584	0.0133***		
	(0.00377)	(0.00607)			(0.00380)	(0.00485)		
Adopter Neighbours								
# Adopter Neighbours			0.0147**	0.0256***			0.00779*	0.0241***
			(0.00606)	(0.00897)			(0.00450)	(0.00583)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1337	1337	1337	1337	1381	1381	1381	1381

Notes: Constant and treatment levels omitted from output. (M): Midline, (E): Endline.
Standard errors clustered at the village level
$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

We find that having an additional plot neighbour does not increase the likelihood of adopting TCB in the control group, whilst finding some evidence of there being an effect in the spillover group. Importantly, we find that having neighbours who adopted TCB is a significant effect in both of these groups. Our results strongly corroborate the evidence we have found with respect to the importance of being connected with adopters. Having more neighbours is has a much greater effect

[29] A caveat: This paper discusses not only technology adoption decisions but also optimal values of farming inputs. Our analysis captures its spirit in the sense that we consider the role of close farming neighbours who share neighbouring farming plots

[30] i.e. neighbours as those living next to you

on the likelihood of TCB take up when those neighbours are themselves adopters of TCB. These results also align with findings in work such as Conley and Udry (2010), which find evidence that the 'observability' of new technologies can enhance social network effects.

In summary, we provide evidence that the key mechanism underlying the association between social connections and adoption outcomes are links with farmers that adopt TCB. This mechanism allows us to connect two, seemingly contradictory pieces of evidence: that connections with treated farmers increase the likelihood that a spillover farmer adopts TCB, and that there are comparable rates of adoption amongst pure control farmers. Firstly, we find that there is evidence of a sort of diffusion effect present within the control villages: adoption at endline is positively associated with links with farmers who adopted at midline. We then test for the presence of this effect within treatment villages. We find that spillover farmers are significantly more likely to adopt TCB if they are linked with farmers who have adopted previously. Moreover, we provide evidence the magnitude of this effect is the same across both control and spillover groups. This suggests that social networks facilitate treatment effects insofar as treatment has a direct affect on farmers who attend. Finally, we buttress our earlier results by showing that there is a low likelihood that social networks are transmitting negative information from non-adopters, and that farming neighbours are much more likely to influence adoption if they are adopters themselves.

6 Conclusion

We examined the relationship between social networks and the adoption of a banana cultivation technology, TCB, in Kirinyaga County, Central Kenya. Our work differs from prior research in that it assesses the role of social networks in facilitating the diffusion of technology adoption decisions in the case where treatment is randomly allocated as opposed to targeted to particular individuals within a network. To the best of our knowledge, this is the first study that has quantified social network effects in settings where different proportions of network members have been randomly allocated to technology information sessions.

Utilising a cluster RCT design with randomisation levels at both the village and farmer level, we exploit the fact that only a proportion of households in treatment villages were invited to training programs in order to identify the role of social connections between treated and untreated farmers on the adoption outcomes of the untreated. The treatment - a training program for TCB cultivation methods - is shown to affect not only those treated directly, but also their network links. We find that an untreated farmer in a treated village is about 2.4 percentage points more likely to adopt the technology if they are connected with an additional treated farmer.

In order to address a key feature of the data: that control and spillover farmers have similar rates of adoption by endline, we examined some potential candidate mechanisms. Our results here are telling. We find that there is strong evidence that links to TCB adopters are a significant predictor of eventual adoption for both control and spillover farmers. Moreover, we find that the size of this effect is the same amongst both control and spillover farmers. This would not only explain high levels of adoption in the control group, but also sheds light onto the mechanism through which treatment effects diffuse through the network. Owing to budget constraints, we only collected network and TCB adoption data at midline and hence are not able to verify whether high levels of baseline adoption are driving the high levels of adoption in the control group.

It would appear that the efficacy of the intervention is enhanced by social networks insofar as they directly increase the likelihood that treated farmers adopt TCB. These farmers who adopt are connected with untreated farmers who, we have shown, are then more likely to adopt themselves. This would suggest that the social network's role in diffusing technology is independent of whether or not network members reside in treatment villages.

The implication here accords which much of the literature discussed in Section 2 in the sense that policy makers should spend most of their attention ensuring those who treated actually adopt the technology, rather than attempting to spread awareness by broadly spreading information to proportions of network members. The evidence from our study does not necessarily support the implementation of targeted interventions over those where treatment is randomly allocated, however it does suggest that more attention should be paid to ensuring those who are treated actually end up adopting the technology – presumably at the cost of treating a smaller sample of farmers. An example of this are the 'goal-setting' interventions discussed in Chowdhury et al. (2019).[31] Future research may investigate how interventions that are more directly aimed at ensuring treated farmers adopt TCB (as opposed to just instructing on the use of technology) are enhanced by the social connections of whom those farmers are connected with.

Future research could consider deeper mechanisms that explain why links with adopters increase the likelihood that a farmers adopts technology. It is not clear exactly what drives these effects, for example, it may be that adopter links are more effective when farmers communicate with a higher frequency or share common social connections. We find some preliminary evidence that sharing neighbouring plots could play a role. One could also use a structural model in the spirit of Banerjee et al. (2013) in order to formally model a distinction between the popularity and diffusion effects discussed in Section 4.1.2.

[31] These goal setting interventions were aimed at addressing procrastination as a barrier for adoption. Future interventions need not necessarily be targeted.

References

Akbarpour, M., Malladi, S., and Saberi, A. (2018). Just a few seeds more: value of network information for diffusion. *Available at SSRN 3062830*.

Alatas, V., Chandrasekhar, A. G., Mobius, M., Olken, B. A., and Paladines, C. (2019). When Celebrities Speak: A Nationwide Twitter Experiment Promoting Vaccination In Indonesia. Technical report, National Bureau of Economic Research.

Bandiera, O. and Rasul, I. (2006). Social networks and technology adoption in Northern Mozambique. *Economic Journal*.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2013). The diffusion of microfinance. *Science*.

Banerjee, A. V., Breza, E., Chandrasekhar, A. G., and Golub, B. (2018). When Less is More: Experimental Evidence on Information Delivery During India's Demonetization.

Beaman, L. A., BenYishay, A., Magruder, J., and Mobarak, A. M. (2018). Can Network Theory-Based Targeting Increase Technology Adoption? *SSRN Electronic Journal*.

BenYishay, A. and Mobarak, A. M. (2018). Social learning and incentives for experimentation and communication. *The Review of Economic Studies*, 86(3):976–1009.

Cai, J., De Janvry, A., and Sadoulet, E. (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics*, 7(2):81–108.

Chandrasekhar, A. (2016). *Econometrics of Network Formation*.

Chandrasekhar, A. and Lewis, R. (2011). Econometrics of sampled networks. *Unpublished manuscript, MIT.[422]*.

Chowdhury, S., Kipchumba, E., Mariara, J., Murigi, M., Nganga, M., Sharma, U., and Sulaiman, M. (2019). Information, procrastination and neighbours' actions: Experimental evidence of adopting improved banana variety in Kenya. *Unpublished Manuscript*.

Cilliers, J., Donnenfeld, Z., Kwasi, S., Shah, S. S., and Welborn, L. (2018). Shaping the future-strategies for sustainable development in kenya. *ISS East Africa Report*, 2018(18):1–32.

Conley, T. G. and Udry, C. R. (2010). Learning about a new technology: Pineapple in Ghana. *American Economic Review*.

Crane-Droesch, A. (2018). Technology diffusion, outcome variability, and social learning: Evidence from a field experiment in Kenya. *American Journal of Agricultural Economics*.

Ewert, U. C. and Sunder, M. (2011). Trading networks, monopoly, and economic development in medieval northern europe: an agent-based simulation of early hanseatic trade. In *9th European historical economics society conference, Dublin*.

Fafchamps, M. and Gubert, F. (2007). The formation of risk sharing networks. *Journal of Development Economics*.

Fafchamps, M. and Quinn, S. (2013). Social networks and business practices: Evidence from a randomized experiment with microentrepreneurs. Technical report, Working paper.

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

Islam, A., Ushchev, P., Zenou, Y., and Zhang, X. (2018). The value of information in technology adoption: Theory and evidence from bangladesh.

Iyer, S. and Weeks, M. (2009). Social interactions, ethnicity and fertility in Kenya.

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

Jackson, M. O. (2011). An overview of social networks and economic applications. In *Handbook of social economics*, volume 1, pages 511–585. Elsevier.

Jackson, M. O., Rogers, B. W., and Zenou, Y. (2017). The economic consequences of social-network structure. *Journal of Economic Literature*, 55(1):49–95.

Maertens, A. and Barrett, C. B. (2012). Measuring social networks' effects on agricultural technology adoption. *American Journal of Agricultural Economics*, 95(2):353–359.

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

Mbugua, M., Nzuma, J., and Muange, E. (2019). Social networks and ex post risk management among smallholder farmers in kenya. *Development Studies Research*, 6(1):30–39.

Miguel, E. and Kremer, M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*.

Miller, G. and Mobarak, A. M. (2014). Learning About New Technologies Through Social Networks: Experimental Evidence on Nontraditional Stoves in Bangladesh. *Marketing Science*.

World Bank, . (2017). *Africa's Pulse, No. 16, October 2017*. World Bank Group.

7 Appendix

A.1 Figures

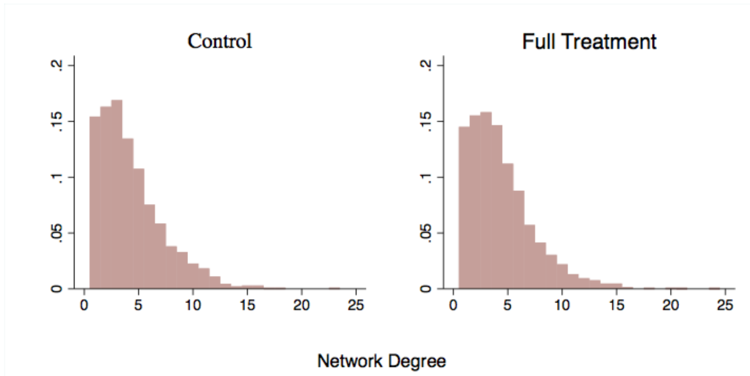

Figure A1: Degree Distributions: Control vs. Treatment

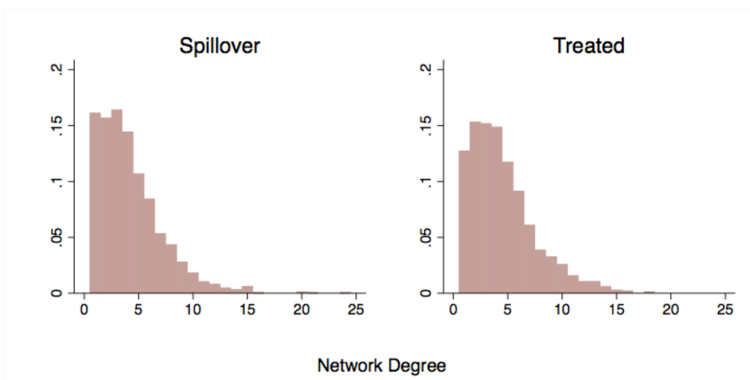

Figure A2: Degree Distributions: Spillover vs. Treated

A.2 Appendix to Experimental Design

A.2.1 Network question allocation method

The process which determined which households a responding household was queried is best explained by example. Suppose we sample a set of households $\{a, b, c\}$ and we are interested in characterising the social connections between these 4 households. We define a social connection based on survey data which we have obtained, and we are interested only in whether or not 2 households are connected or not, that is social connection is represented by 0 if the pair is not

Section 18: Social Network

Connection between treatment and control groups. For the treatment farmers, ask about all the control farmers. For control group farmers, ask about the treatment group. Not applicable for pure control villages.

Select the farmers of the village (use Household ID)	Do you know [NAME]? 1=Yes 0=No If No→next name	How is [NAME] related to you? 1= Relative 2=Friend 3=Acquintance 4=Other	How long have you known [name]? 1=Less than a year 2=2-3 years 3= More than 3 years	How often do you talk to [Name]? 1=Everyday 2=Weekly 3=Monthly 4=Rarely	Is [NAME] in any farmer group with you? 1=Yes 0=No	Do you have a plot that is neighbouring a cultivation plot of [NAME]? 1=Yes 0=No 99=DK	(if yes) Is banana grown on any of these plots? 1=Yes 0=No	Interactions in the last 12 months		
								Have you borrowed from or lent money to [NAME]? 1=Yes 0=No	Have you discussed about banana cultivation practices with [NAME]? 1=Yes 0=No	Have you done any banana sales together with [NAME]? 1=Yes 0=No
ID	1	2	2a	3	4	5	5a	6	7	8

(Thank the respondent and collect GPS coordinates of the interview location)

Figure A3: Social network questions asked by EADMA representatives

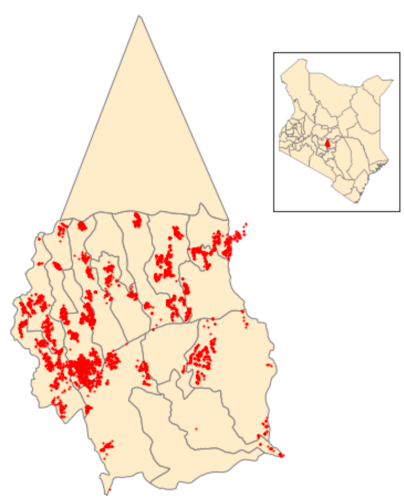

Figure A4: Sampling Locations in Kirinyaga County, Kenya. Map of Kenya inset.

linked and 1 if the pair is. Now, there are $\frac{(3)(3-1)}{2} = 6$ total combinations of pairs that can be defined within this set of households:

$$(a, b), (b, a), (a, c), (c, a), (b, c), (c, b) \tag{8}$$

If we assume that the social network is 'undirected' (i.e. $i \overset{link}{\rightarrow} j \Leftrightarrow j \overset{link}{\rightarrow} i$), we may observe the full network (the existence of a link or non-link between every pair within the network) by restricting our attention to the *unique* pairs of households within the network: $(a, b), (a, c), (b, c)$. Our sampling methodology relies on this logic. The pseudo code for the methods is as follows:

Step 1: Obtain all pairs of households within a sample

Step 2: Sort this list such that (a, b) follows (b, a) and so forth for all of the pairs as demonstrated in (8)

Step 3: Obtain all the unique pairs whilst maintaining the order imposed by the sorting in Step 2

Step 4: For each unique pair, ask the first entry the set of social network questions regarding the second entry e.g. if (a, b) is one of the unique pairs obtained in Step 3, then ask the respondent of household a questions regarding household b

Sampling conducted this way ensures that the survey will in theory (assuming no attrition or non-response), elucidate information regarding the social connections between all nodes (households) within the village-level network. The survey intended to ask each household within the sample about $(n-1)/2$ other households, which households they were asked about was determined by the process stated above.

A.3 Preliminary Results

In this first section we outline some preliminary results which a) establish the direct effect of the interventions themselves and b) provide some preliminary evidence for the existence of the association between social network characteristics and technology adoption outcomes.

A.3.1 Direct Effects of Treatment

We provide evidence that the intervention was effective on households which were actually invited to treatment. Since we control for whether a household was invited to the training programs

Group	***Controls***	***N***
Control		1371
	✓	1337
Spillover		1419
	✓	1381
Treated		1310
	✓	1280
Full		4101
	✓	3998

Notes: N: Number of observations. Entries associated with controls are number of observations for which a full set of controls exists

Table A9: Number of Observations by Group

Table A10: Attrition and Degree

	(1)	(2)
Degree	0.000568*	0.000436
	(0.000317)	(0.000399)
Spillover=1 × Degree		-0.000228
		(0.000435)
Control=1 × Degree		0.000637
		(0.000437)
Observations	4101	4101

Notes: Constant omitted from output. Standard errors clustered at village level. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

rather than considering which households actually attended, we report 'Intention To Treat' (ITT) estimates.[32]

Table A12 reports ITT estimates from the following specification:

$$Adopted_{itv} = \beta_0 + \beta_1 \mathbf{1}\{Treated\}_{iv} + \beta_2 \big[\mathbf{1}\{Treated\}_{iv} \times Intensity_v\big] + \gamma' \mathbf{x}_{iv} + u_{iv} \tag{9}$$

Where $Adopted_{itv}$ equals 1 if individual i in village v adopts TCB by time t. $\mathbf{1}\{Treated\}_{iv}$ is an indicator variable equalling 1 if the household was invited to the EADMA training programs and $Intensity_v$ is a set of indicators representing treatment intensity and $\mathbf{x}_{iv}$ is a set of controls. We find strong evidence that the treatment was effective on households that were treated. Compared to control households, treated households are 10.6 percentage points more likely to adopt TCB by endline. These results echo findings outlined in Chowdhury et al. (2019) which provides a comprehensive program evaluation of the training program effects.

A.3.2 The Association Between Social Networks and Adoption Outcomes

First, we present results that evidence the correlation between network variables and adoption outcomes. We estimate:

$$Adoption_{iv} = \beta_0 + \beta_1 Network_{iv} + u_{iv} \tag{10}$$

$Adoption_{iv}$ equals 1 if household i in village v adopted TCB by either midline or endline and 0 otherwise and $Network_{iv}$ is the network characteristic of interest.

We use 2 different network characteristics: degree and eigenvector centrality. Results are reported in Table A11. Columns 1 and 4 suggest that a household's degree is positively associated with the propensity that it adopts TCB by both midline and endline. Although small, these results as suggest that households who are connected with a greater number of village neighbours are more likely to adopt the TCB technology irrespective of whether or not they are treated. We test these measures jointly in Table A13 and find strong evidence of their joint significance.

As discussed in the literature review, much of the work that has utilised social networks in development settings has focused on network targeting. A popular tool for measuring which nodes

[32] ITT: 'Intention To Treat'. There are moderate levels of treatment compliance within villages with about 68% of invitees attending training program, more details can be found in Chowdhury et al. (2019)

Table A11: Network Characteristics on Adoption Outcomes

	Individual Regressors				Both Regressors	
	(M)	(M)	(E)	(E)	(M)	(E)
Degree	0.00750***		0.0130***		0.00742***	0.0128***
	(0.00219)		(0.00380)		(0.00224)	(0.00380)
Eigenvector Centrality		-0.754		-1.757	-0.472	-1.271
		(1.266)		(2.071)	(1.268)	(2.003)
Observations	4101	4101	4101	4101	4101	4101

Constant omitted from output. Standard errors clustered at the village level
*p ¡0.10,**$p < 0.05$,*** $p < 0.01$

may be influential within a network is 'eigenvector centrality'[33] (see Section 3.5). Studies such as Banerjee et al. (2013) for example, provide evidence that village-level adoption rates are positively correlated with the average eigenvector centrality of treated individuals. We actually find evidence to the contrary; Table **??** illustrates that the average eigenvector centrality of treated individuals is actually negatively associated with village-level adoption outcomes. Nonetheless the results shown in Table A11 indicate that our key variable of interest, network degree is positively associated with the adoption rates.

Table A12: Direct Effect of Treatment

	Midline		Endline	
	(1)	(2)	(3)	(4)
Treatment HH=1	0.0328**	0.0105	0.106***	0.00512
	(0.0154)	(0.0228)	(0.0323)	(0.0555)
Treatment HH=1 × Intensity=20		0.130**		0.143*
		(0.0628)		(0.0848)
Treatment HH=1 × Intensity=40		0.0271		0.0825
		(0.0313)		(0.0806)
Treatment HH=1 × Intensity=60		0.0424		0.239***
		(0.0329)		(0.0663)
Control (Mean)	0.117***	0.108***	0.273***	0.271***
	(0.0109)	(0.0140)	(0.0176)	(0.0273)
Observations	4101	4101	4101	4101

Notes: Standard errors clustered at village level.
$^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

[33]Eigenvector Centrality can be defined implicitly by: $\mathbf{Ax} = \lambda\mathbf{x}$ where $\mathbf{x}$ is an n-dimensional vector of eigenvector centrality measures, $\mathbf{A}$ is the adjacency matrix described in Section 3.1.2 and λ is the eigenvector that solves the equation.

Table A13: Joint Significance of Network Characteristics

	(1) Adopted Midline	(2) Adopted Endline
Degree	0.00797	0.0130
	(3.58)	(3.44)
Eigen. Centrality	-0.442	-1.120
	(-0.35)	(-0.56)
Clustering	-0.0394	0.0156
	(-1.61)	(0.44)
F	5.601	4.274

Model 1 p-value: 0.0015, Model 2 p-value: 0.0072

A.4 Appendix to Main Results

A.4.1 Probit Estimation

Tables A14 and A15 replicate the results obtained from equations (1) and (2) but with Probit specifications. Formally, the results from Tables A14 and A15 are results from the estimation of:

$$\begin{aligned} Pr(Adoption^G_{iv} = 1|\mathbf{X}) = \Phi\Big(\beta_0 + &\beta_1 Degree_{iv} + \beta_2 Treatment\ Level_v \\ &+ \beta_3[Degree_{iv} \times Treatment\ Level_v] + \gamma'\mathbf{x}_{iv}\Big) \end{aligned} \tag{11}$$

$$\begin{aligned} Pr(Adoption^S_{iv} = 1|\mathbf{X}) = \Phi\Big(\beta_0 + &\beta_1 Degree^T_{iv} + \beta_2 Treatment\ Level_v \\ &+ \beta_3[Degree^T_{iv} \times Treatment\ Level_v] + \gamma'\mathbf{x}_{iv}\Big) \end{aligned} \tag{12}$$

A.4.2 Alternative Network Measures: Link Density

Recall, our main results utilise ‘degree’ which is a count of how many links a particular households (see Section 3.5). As a further robustness check, we test whether our main results are sensitive to the manner in which we define our network variables of interest. In order to do this, we utilise a different, albeit similar measure called ‘link density’. Link density is similar to degree however it

Table A14: Probit: Degree on HH Adoption By Group

	Control		Treated				Spillover			
	(M)	(E)	(M)	(E)	(M)	(E)	(M)	(E)	(M)	(E)
Degree	0.0521***	0.0356*	0.0228*	0.0212	0.00321	0.0128	0.0289**	0.0449***	0.0204	0.0414***
	(0.0193)	(0.0200)	(0.0137)	(0.0160)	(0.0335)	(0.0476)	(0.0145)	(0.0132)	(0.0204)	(0.0156)
Int.=40 × Degree					0.0353	-0.00790			0.0226	0.0116
					(0.0453)	(0.0593)			(0.0357)	(0.0347)
Int.=60 × Degree					-0.0145	0.0352			0.00578	0.0126
					(0.0428)	(0.0552)			(0.0432)	(0.0365)
Int.=80 × Degree					0.0463	-0.00317			0.0184	-0.0197
					(0.0373)	(0.0539)			(0.0473)	(0.0378)
Controls	No	No	No	No	No	No	No	No	No	No
Observations	1371	1371	1310	1310	1310	1310	1420	1420	1420	1420

Notes: Constant and treatment levels omitted from output. (M): Midline, (E): Endline.
Standard errors clustered at the village level
$^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

Table A15: Probit: Links with Treated HHs on Adoption Outcomes

	Spillover				Control (Pseudo-Treatment)			
	(1) (M)	(2) (E)	(3) (M)	(4) (E)	(5) (M)	(6) (E)	(7) (M)	(8) (E)
Links (Treat. HHs)	0.0282	0.0708***	0.0149	0.0299				
	(0.0247)	(0.0203)	(0.0360)	(0.0315)				
Int.=40 × Links (Treat. HHs)			0.100	0.103**				
			(0.0619)	(0.0424)				
Int.=60 × Links (Treat. HHs)			-0.0323	0.0696				
			(0.0603)	(0.0562)				
Int.=80 × Links (Treat. HHs)			-0.0582	0.0120				
			(0.0705)	(0.0665)				
Links (Pseudo Treat. HHs)					0.000707	0.0297	0.0286	-0.0321
					(0.0294)	(0.0319)	(0.0390)	(0.0397)
Pseudo Int.=40 × Links (Pseudo Treat. HHs)							-0.00244	0.123**
							(0.0594)	(0.0616)
Pseudo Int.=60 × Links (Pseudo Treat. HHs)							-0.0815	0.0868
							(0.121)	(0.0754)
Pseudo Int.=80 × Links (Pseudo Treat. HHs)							-0.482**	0.0225
							(0.195)	(0.115)
Controls	No	No	No	No	No	No	No	No
Observations	1420	1420	1420	1420	686	686	686	686

Constant omitted from output. Standard errors clustered at village level
M: Midline, E: Endline. Controls in all regressions.
$^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

is a proportion rather than a count. It is used within various parts of the literature (e.g. Cai et al. (2015)). More concretely:

1. *Density:* Proportion of households the household i is connected with.

$$Density_{iv} = \frac{\#Links\ within\ village_{iv}}{\#HHs_v}$$

2. *Density of Treated HHs:* Proportion of households the household i is connected with. This measure is only observable for spillover sample households.

$$Density^T_{iv} = \frac{\#Links\ with\ Treated\ HHs_{iv}}{\#Treated\ HHs_v}$$

We utilise these measures and re-estimate equations (1) and (2). We report results in Tables **??** and **??**. We interpret these coefficients slightly differently.For example, within the spillover sample, a unit increase in link density increases the probability that a household takes up TCB by endline by 68 percentage points (Table **??**; Column 6, Column 8). We can interpret this association as indicating that having an additional social connection being invited to a training program increases the likelihood that a household adopts TCB by $0.682 \times 0.02 = 1.36\%$. These results corroborate our earlier findings, the magnitude of the effects we observe is also comparable.

A.5 Psuedo Treatment Method

The aim here was to allocate pure control households into spillover and treated groups. We call these groups the 'pseudo-spillover' and 'pseudo-treated'[34] groups respectively. The idea then, was to test whether links between these households would be predictive for adoption outcomes. Since we find a significant effect for these links in the true spillover sample (see Table 4), we construct these pseudo-groups in order to create a counter factual. Our hypothesis is that if diffusion does not occur, then links between these randomly allocated households, in the presence of no treatment, should not affect adoption outcomes.

The pseudo code for the pseudo treatment algorithm is as follows.

Step 1: Set random number generator seed

Step 2: Randomly sort villages into 20%, 40%, 60% and 80% treatment groups

[34]The 'pseudo' comes from the fact that these households of course, do not actually receive any treatment directly

Step 3: Within each village v, generate a random number for each household i and then sort the n_v observations in ascending order

Step 4: Allocate a number of households (which corresponds with treatment level randomised in Step 2) to pseudo-treated group, allocate the rest to the pseudo-spillover

Step 5: Define an adjacency matrix for each village

Step 6: Calculate degree amongst pseudo-treated and pseudo-spillover households for each village

Step 7: Estimate equation (3) and store coefficients

The results reported in Table 4 are clearly sensitive to the seed that is chosen in Step 1. In order to ensure robustness we repeat steps 1-7 30 times[35] and report the amount of times we reject the coefficient β_1 in equation 3. This coefficient captures the association between pseudo degree and adoption of TCB at endline. We found that the coefficient was significant in only $\frac{2}{30} \approx 0.06$ cases.

Table A16: Relationship Between Links With Treated HHs and Treatment Intensity

	Spillover		Pseudo Spillover	
	(1) Degree	(2) Density	(3) Degree	(4) Density
Levels of Treatment Intensity				
Intensity=40	0.647**	-0.0111	0.232	0.00497
	(0.308)	(0.00930)	(0.373)	(0.00776)
Intensity=60	0.876**	-0.0186**	0.0453	0.00240
	(0.371)	(0.00909)	(0.243)	(0.00489)
Intensity=80	0.0414	-0.0332***	0.0361	0.00545
	(0.319)	(0.00873)	(0.341)	(0.00653)
Controls	Yes	Yes	Yes	Yes
Observations	1381	1381	1352	1352

Constant omitted from output. Standard errors clustered at village level
We report pseudo network characteristics and intensity for pseudo sample
$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

[35] Due to time constraints we could only run this algorithm 30 times, although we intend on running it more times in the future