Network-Based Targeting with Heterogeneous Agents for Improving Technology Adoption

Preliminary and Incomplete - Please do not Cite without Permission

Aranya Chakraborty*

August 2022

Abstract

How do we use existing social ties to improve the adoption of a new technology? I explore network-based targeting when the benefits from the technology vary at the household level, with this heterogeneity in benefits affecting the diffusion of information. I develop a theoretical framework where initially uninformed agents engage in DeGroot learning to decide whether or not to get fully informed about a new technology. Conditional on being fully informed, they then decide whether or not to adopt the technology. The model predicts the possibility of low information equilibria where nobody will adopt the new technology even if it is the efficient choice for some of them, and targeting is needed. My simulations suggest that the optimal targeting strategy in such a scenario relies on the underlying heterogeneity in the population. If heterogeneity is low in the benefits from the technology, targeting based on centrality works well. However, if the population is highly heterogeneous, centrality-based targeting fails in reaching the population of interest. In such a scenario, targeting based on the probability of adoption works better if the network is highly assortative in terms of characteristics determining the heterogeneity. I test these predictions using data from Malawi and provide evidence supporting my theoretical model. I argue that population heterogeneity in benefits from a technology matters for the success or failure of alternative targeting strategies that promote that technology.

JEL Codes: D83, O13, O33, Q16

Keywords: Targeting, Social Network, Technology Adoption, Agriculture

---

*Department of Economics, McGill University; Canada. E-mail: aranya.chakraborty@mail.mcgill.ca.
I am grateful to Sonia Laszlo and Francesco Amodio for their continuous guidance and support throughout the project. I also thank Matthieu Chemin, Franque Grimard, Leonie Baumann, Mallesh Pai and Adam Aberra for their helpful comments. I would also like to extend my gratitude to the participants at the McGill-CIREQ Lunch Seminars, the 56th Annual Conference of the Canadian Economics Association (CEA), and the 17th CIREQ Ph.D. Students' Conference for their generous feedback. I gratefully acknowledge the support from all the faculty members of the Department of Economics at McGill University. I am indebted to OPENICPSR and Harvard Dataverse for access to the datasets used in this study. All views expressed in this work are my own.

1 Introduction

Technology adoption in agriculture is a driving force of economic growth through its effect on structural transformation (Bustos et al., 2016). However, the adoption of modern technologies has been low in developing regions, especially in Sub-Saharan Africa (Bold et al., 2017). Information constraints are one of the key reasons behind such phenomenon (Magruder, 2018). How do we use existing social ties to improve the adoption of a new technology? The literature argues that the answer depends on the underlying diffusion process. If information diffuses only if a certain threshold of each agent's connections is informed, targeting based on existing social ties may be required for widespread adoption. In such a scenario, the literature recommends targeting agents central to the network (Beaman et al., 2021a). The recommendation, however, is based on the underlying assumption that the diffusion only depends on the agents' positions in the network. What happens if the agents differ in terms of other characteristics that affect the diffusion process?

This paper investigates network-based targeting strategies for improving technology adoption. In particular, I focus on the situation where the new technology can be more beneficial to some agents than others, with this heterogeneity in benefits affecting the diffusion of information. The benefits can vary across agents due to several possible reasons. The agents can differ in terms of their education, skills, and ability affecting how much they can learn about a new technology and use it in practice. They can also vary in terms of other characteristics, e.g., land quality (for agriculture), size of operation (for both farm and firm households), access to infrastructure (such as road and irrigation facilities), and access to other technologies. For my purpose, I consider heterogeneity in benefits to reflect the existing network structure driven by agent sorting according to their observable and unobservable characteristics. I explore whether the optimal network-based targeting strategies vary with the extent of heterogeneity within the network. More specifically, I concentrate on the relative performance of two targeting strategies: targeting based on centrality and targeting based on probability of adoption.

I develop a theoretical framework where economic agents participate in a two-stage decision process. In the first stage, uninformed agents decide whether or not to get fully informed about a new technology. Since information is costly, the agents engage in DeGroot learning to make this decision.[1] In the second stage, fully informed agents decide whether or not to adopt the technology. This framework helps me formalize a scenario where pessimism regarding the prospect of a new technology will lead to low

[1]DeGroot learning refers to a social learning process whereby agents form beliefs/ opinions as a weighted average of the beliefs/ opinions of people they are linked to (including themselves). Here the weights correspond to how much the agents are influenced by one another. It is a heuristic, as agents do not account for the interdependence of beliefs between each of the people they are connected to (Barnett-Howell and Mobarak, 2021). Chapter 8.3 of Jackson (2010) contains more information on this type of learning.

adoption, even if it is efficient for many agents to adopt.

Based on my theoretical model, I use simulations to evaluate the relative importance of different targeting strategies and to generate testable hypotheses.[2] I test these predictions by combining two different data sources from Malawi. The first one is the replication data (Beaman et al., 2021b) from a randomized controlled trial (RCT) conducted by Beaman, BenYishay, Magruder, and Mobarak (2021a) (henceforth, BBMM). The second dataset is the Agricultural Extension Services and Technology Adoption Survey (henceforth, AESTAS) data (IFPRI, 2021a,b) collected by International Food Policy Research Institute (IFPRI). One of the reasons existing studies made simplifying assumptions on the structure of heterogeneity in the population is the difficulty in observing heterogeneity in benefits beforehand. As the benefits are only realized after adoption, they cannot be factored into the targeting strategies. I attempt to solve this issue by using AESTAS data to estimate adoption conditional on observable demographics. This way I can categorize the population in terms of their propensity to adopt a new technology. I calculate households' probability of adoption in the BBMM data using estimates from the AESTAS sample. BBMM data is used as their experiment relies on exploiting the centrality of seeds to improve the adoption of a technology suitable for my analysis, thus including all other information that I need. I exploit both the village-level and experimental variations in the BBMM data to test my hypotheses.

My simulations indicate that the relative performance of different targeting strategies depends on the degree of heterogeneity in a network. Centrality-based targeting strategies should be less effective in settings where the agents vary significantly in terms of their true benefits from adopting a technology. In such settings, targeting based on the likelihood of adoption should perform better if the network is highly assortative in terms of characteristics determining the benefits. The intuition behind such a result lies in the characteristics of the central seeds in a network.[3] Central seeds are, by definition, the most well-connected people in a network. Thus, selecting them would maximize the diffusion if diffusion depends only on the agents' positions in the network. If agents vary in terms of other characteristics that affect diffusion, we need to consider this heterogeneity for effective diffusion. Centrality-based targeting fails to consider this heterogeneity. In an assortative network, central seeds also represent the average network characteristics. In a setting where a new technology applies to only a certain sub-section of the population, targeting based on centrality becomes more likely to fail in reaching the population of interest. Targeting the population of interest works

[2]The use of simulations is not new to the network literature. For example, Bala and Goyal (1998) uses simulations to generate spatial and temporal patterns of adoption when individuals learn from their neighbors; Acemoglu et al. (2011) uses simulations to show that innovations might spread further across networks with a smaller degree of clustering. Similar to Beaman et al. (2021a), I use them to understand the effectiveness of targeting strategies a few periods down the line.

[3]In network literature, information entry points are termed *seeds*.

better in such a scenario.

Reduced form results show evidence in favor of my hypothesis. Exploring village-level variations in the BBMM data, I show that the positive effect of seeds' centrality on the adoption of pit planting decrease with an increase in village-level heterogeneity in terms of probability of adoption. Simultaneously, the negative effect of seeds' probability of adoption decreases with increased village-level heterogeneity. Weaker, but similar results are found when I shift my focus to exploring experimental variations.

My study makes three contributions to the existing literature. First, I provide evidence (both theoretical and empirical) that the success of network-based targeting strategies depends on population-level heterogeneity. Diffusion of information via networks is the key to increasing technology adoption (Besley and Case, 1993; Foster and Rosenzweig, 1995; Conley and Udry, 2010; Krishnan and Patnam, 2013). In recent years, there have been several studies focusing on the role of networks in the diffusion of technologies.[4] A growing proportion of these studies explore the most effective way to use social networks to improve technology adoption (e.g., Banerjee et al., 2013; BenYishay and Mobarak, 2018). A few of these studies explore the role of the underlying diffusion process in designing the most effective targeting policies (e.g., Beaman et al., 2021a; Akbarpour et al., 2021). However, these studies assume existing network ties to be the only factor characterizing diffusion. Thus, for diffusion, households are assumed to be homogeneous in terms of other characteristics. In the current study, I consider the population to be heterogeneous in terms of the benefits they get from the new technology, with this heterogeneity having a direct effect on the effectiveness of targeting strategies. In such a scenario, I show evidence that optimal targeting strategies may differ from the ones prescribed in the existing literature. In particular, the effectiveness of a targeting policy will vary depending on population-level heterogeneity in terms of the benefits of the new technology. Considering population-level heterogeneity in social learning itself is not new (e.g., Munshi, 2004; Bandiera and Rasul, 2006; Conley and Udry, 2010).[5] However, to the best of my knowledge, the current study is the first to consider the consequences of population-level heterogeneity on network-based targeting strategies.

Second, my theoretical framework helps formalize the scenario where agents learn from their network about a technology that is more beneficial to some of them than others. Existing studies consider technologies to be equally beneficial to everyone. The adoption may still differ due to heterogeneity in costs. But these heterogeneous costs are assumed to be known by the agents and thus do not require learning.[6] Thus,

[4]See Cheng (2021) for a review of the existing literature.

[5]Using the data from Indian Green Revolution, Munshi (2004) finds that information flows are weaker for rice growers than wheat growers as rice-growing regions are more heterogeneous. Bandiera and Rasul (2006) observe network effects on technology adoption to vary based on the number of adopters in the network for sunflower production in Mozambique. Conley and Udry (2010) finds that only novice farmers learn from their veteran neighbors about the use of fertilizers for pineapple production in Ghana.

[6]Even if the heterogeneous costs are unknown to the agents, there is no possibility of learning from

simplifying assumptions are made such that the learning involves the characteristics that are similar for all the agents and not the characteristics that differentiate them. This assumption helps us to focus on a problem where the agents are collectively trying to uncover some hidden characteristics of interest (e.g., in the theoretical models of Acemoglu et al., 2008 and Golub and Jackson, 2010). More importantly, a consequence of this assumption is that the diffusion of knowledge regarding the technology depends only on the agent-level heterogeneity in network ties. In many scenarios, however, agents face heterogeneous benefits in adopting a new technology (Suri, 2011). For example, the performance of some agricultural practices may depend on the land quality.[7] Thus, the benefits of some technologies may vary depending on the agent-specific characteristics. The consequences of this heterogeneity on the diffusion of knowledge have not previously been studied in the existing literature.

Finally, I provide policy directions for network-based targeting when the population is heterogeneous. In particular, I argue in favor of targeting early adopter households when the heterogeneity is high and if the network is highly assortative in terms of characteristics determining the heterogeneity.[8] On the contrary, I argue in favor of targeting central households when the heterogeneity is low. This policy recommendation contributes to the literature that focuses on understanding the characteristics and impact of opinion leaders in diffusing new knowledge. In this literature, studies like Maertens (2017) and Miller and Mobarak (2015) show that learning is more effective when the opinion leaders are in some way *superior* than their followers. On the other hand, BenYishay and Mobarak (2018) show that communicators who share a group identity with the farmers or face comparable agricultural conditions, do a better job at convincing farmers to adopt a new technology. Feder and Savastano (2006) takes a middle ground in arguing that the most effective opinion leaders are superior to their followers, but not excessively so. My study contributes to this debate from a network-based targeting perspective.

The remainder of this article is organized as follows. In Section 2, I present the theoretical framework of my analysis. Section 3 presents the simulations that help me form the main hypotheses for this study. In Section 4, I discuss the hypotheses, my empirical strategy for testing them, and the data I use in the process. Section 5 presents and discusses my empirical results. Finally, in Section 6, I summarize my findings and

the network as these costs are not correlated according to the network structure. There can still be a possibility of learning by doing.

[7]In Munshi (2004), the adoption of new rice varieties is sensitive to growing conditions. Tjernström (2017) shows that soil quality heterogeneity affects farmers' ability to learn from their peer's experimentation with the new technology. Pit planting studied in BenYishay and Mobarak (2018) and Beaman et al. (2021a) requires flat land.

[8]I define early adopters as the households that are more likely to adopt a new technology given homogeneous cost. This definition is similar to that of natural early adopters in Catalini and Tucker (2017).

make concluding remarks.

2 Theoretical Framework

I consider a choice problem that requires learning in a social network. The problem is that of technology adoption when the agents vary in terms of the benefits they get from a new technology. In particular, the benefits are such that it is optimal to adopt the new technology only for a sub-section of the population. However, the benefits are initially unknown to the agents, who must get informed first before making the adoption decision. As information is costly, agents rely on their social ties to determine whether or not to seek information.

Similar to Golub and Jackson (2010), I consider agents to have an initial opinion and involve in DeGroot learning (developed in DeGroot (1974) and DeMarzo et al. (2003)).[9] I focus on the scenario where the underlying state is time-varying, similar to Acemoglu et al. (2008). Like Banerjee et al. (2021) my model considers both informed and uninformed agents, where agents decide whether to get informed about the new technology.[10] In addition, I consider the possibility that agents are heterogeneous in terms of their distribution of payoffs associated with the new technology.

2.1 The Theoretical Model

Consider a two-stage decision process where households decide whether or not to adopt a new technology. In the first stage, they decide whether or not to learn about the technology. If they learn about the technology, in the second stage, they decide whether to stick to a traditional technology or adopt the new technology.[11]

The traditional technology has a sure payoff of π^T, whereas the new technology provides a payoff of $\pi^N(\omega_{it})$ that depends on the state of the world parameter $\omega_{it} \in \Omega$. The state of the world parameter ω_{it} is drawn independently at each period t according to the true distribution $p_i^*(\omega_{it})$ for household i. Therefore, the draws are not correlated over time within a household and between households.[12] I assume that the true distributions are positively correlated between households according to the existing

[9]DeGroot learning is considered as it is used in all the canonical models of information aggregation in the development literature. There is also empirical evidence in favor of it (see Chandrasekhar et al., 2020).

[10]In Banerjee et al. (2021), uninformed agents have empty beliefs, and informed agents can be partially or fully informed. In contrast, I assume uninformed agents to have an initial opinion (this includes partially informed agents) and informed agents to be fully informed.

[11]This two-step decision process is similar to the model presented in Chandrasekhar et al. (2018). Which also forms the micro-foundation of Beaman et al. (2021a).

[12]The assumption of draws not being correlated over time within a household helps me abstract away from learning-by-doing, where households are observing the draws over time and updating their beliefs accordingly. The assumption of draws not being correlated over time between households constrains the ways the households can learn from each other.

network structure (more details on this below). I also assume that $\forall it$, $\exists \omega_{it}, \omega'_{it} \in \Omega$ such that $\pi^N(\omega_{it}) \geq \pi^T \geq \pi^N(\omega'_{it})$; i.e., for each household i and period t, there exist states of the world such that the payoff from the new technology is higher (lower) than the old technology. Finally, $\exists i, j \in \mathcal{I}$ such that $\int_{\omega_{it} \in \Omega} p_i^*(\omega_{it})\pi^N(\omega_{it}) - c_i \geq \pi^T$ and $\int_{\omega_{jt} \in \Omega} p_j^*(\omega_{jt})\pi^N(\omega_{jt}) - c_j \leq \pi^T$, where $\mathcal{I}$ denote the set of all households and c_i is the cost of new technology for household i. The assumption implies that there is enough heterogeneity in the population such that some households get positive net benefits from adopting the new technology as opposed to the traditional technology, while others do not. This last assumption ensures that the new technology is 'better' for only a fraction of households in the population.

Households can be potentially uninformed about their true distribution for the state of the world. The household i has beliefs $p_{it}(\omega_{it})$ over the distribution of ω_{it} at period t. Every period, an uninformed household has the option to become informed by applying effort $e_{it} \in \{0, 1\}$.[13] Households put effort only once, i.e., if $e_{i\tau} = 1$, $e_{it} = 1 \ \forall t \geq \tau$. If $e_{it} = 1$, the household learns the true distribution $p_i^*(\omega_{it})$ at cost η_i. Households incur the cost of learning only once - the first time they get informed. Households incur no effort cost if $e_{it} = 0$, and they use DeGroot averaging to approximate the true distribution. Let G denote the $n \times n$ weighted and non-negative influence matrix ($n = |\mathcal{I}|$), where $G_{ij} \geq 0$ represents the weight i places on j's opinion (with $\sum_{j \in \mathcal{I}} G_{ij} = 1$ and $G_{ii} \neq 0$). Then $\hat{p}_{it} = \sum_{j \in \mathcal{I}} G_{ij} p_{jt-1}$ denotes household i's approximation based on their aggregation of opinions following the DeGroot averaging.

I assume the networks to be assortative in terms of characteristics that determine the probabilities, i.e., $G_{ij} \neq 0$ if $|p_i^* - p_j^*| < \delta$, where δ is a small number. The rationale behind such an assumption is twofold. First, it is well recognized in the literature that connected agents share similar characteristics (Munshi, 2007). If we focus only on the geographic connections, it is easy to argue that neighboring farmers share comparable soil quality. However, for network connections, the similarity extends beyond geographic characteristics. We expect households to sort according to their socio-economic attributes. Additionally, the *assortative* property is necessary for the social ties to be informative with varying p_i^*s. To demonstrate this, consider Figure 1. Both the panels of this figure present heterogeneous networks, with the colors representing benefits from some technology. For this example, consider red to represent agents with high benefits, blue to represent agents with low benefits, and yellow to represent agents with moderate benefits. Panel A of the figure presents an assortative network. Here uninformed agents can use their social ties to form a belief close to the true type. To see this, consider agent 4, who would benefit highly from the technology.

[13] I make no assumptions on the initial number of informed households. In reality, whether or not a household is informed might depend on their education, skills, and abilities. As I will argue in the next subsection, from a policy perspective, I am interested in the scenario where the majority (if not all) of the households are uninformed about the new technology, to begin with.

If uninformed, she would seek information from agents 3 and 5. Both agents 3 and 5 have higher than average benefits from the technology. Thus, if informed, they can push agent 4 in the right direction because they have the information relevant for agent 4. Contrast this with the network in Panel B, which is a non-assortative network. Here also, agent number 4 would benefit highly from the technology. But, she is connected to agents 1 and 2, both having lower benefits from the technology. Thus, if agent 4 is uninformed and seeks information from her network ties, she will be pushed in the wrong direction as her informed connections would not have the information she requires. Thus, social learning is not possible for the network in Panel B.[14]

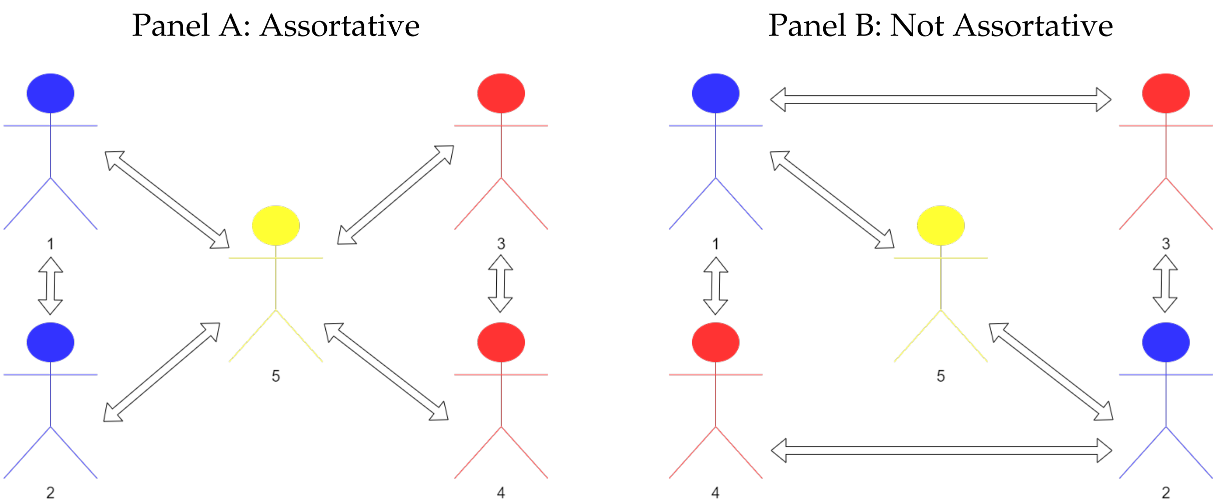


Figure 1: Networks with Heterogeneous Benefits

The belief of household i at period t is determined by the following process:

$$p_{it}(\omega_{it}) = e_{it}(p_i^*(\omega_{it})) + (1 - e_{it})\hat{p}_{it}(\omega_{it}). \tag{1}$$

Thus, uninformed households use DeGroot averaging to approximate the true distribution with the help of their peers. On the other hand, informed households can observe the true distribution and hence do not need to approximate it anymore. Their effort level determines whether they are informed or uninformed. In addition, I assume that households need to be informed before they adopt. As can be seen below, this assumption helps me explicitly capture the point when the households stop seeking information from their peers.

I assume the households to be risk-neutral and myopic. The assumption of risk-neutrality is for simplification purposes only, as it allows us to focus solely on the expected values, without the need to think about the variation around them. As the new technology is assumed to be riskier than the traditional technology here, risk-averse

[14] It is worth noting that if we assume $p_i^* = p^* \ \forall i \in \mathcal{I}$, the network ties become automatically informative.

households may find it less attractive. As such, the net benefit of the new technology would be less than the one perceived by a model where the households are risk neutral. This can easily be accommodated in the current model by dividing the expected payoff of the new technology by its variance. Such an exercise would not change the main results of the model. The assumption of myopic households helps me focus on a static model instead of a dynamic one. More importantly, if the households are not myopic they may wait until their peers get informed before they decide whether or not to get informed themselves. This may lead to a more complicated scenario where everyone is waiting for their peers to get informed first. Such a scenario is beyond the scope of this paper.

Under the above assumptions, a household's adoption decision is a two-step process:

1. First they decide whether or not to get informed, based on the following rule:

$$e_{it} = \begin{cases} 1\ if\ \int_{\omega_{it}\in\Omega} \hat{p}_{it}(\omega_{it})\pi^N(\omega_{it}) - c_i - \pi^T \geq \eta_i \\ 0\ otherwise. \end{cases} \tag{2}$$

   Only uninformed households make this decision.

2. Conditional on being informed, they decide whether or not to adopt the new technology:

$$Adopt_{it} = \begin{cases} 1\ if\ \int_{\omega_{it}\in\Omega} p_i^*(\omega_{it})\pi^N(\omega_{it}) - c_i \geq \pi^T \\ 0\ otherwise. \end{cases} \tag{3}$$

Finally, I assume the following timeline of decision-making:

1. Every period, uninformed household i decides whether or not to get informed. Informed households do not need to get informed as they already know their true probability distribution.

2. To decide, uninformed households collect information on beliefs from their peers $j \in \mathcal{I}$. These beliefs p_{jt-1} were formed in the last period (some informed, some uninformed). Using their peers' beliefs from the last period, the uninformed household i uses DeGroot averaging to calculate $\hat{p}_{it} = \sum_{j\in\mathcal{I}} G_{ij}p_{jt-1}$.

3. Based on $\hat{p}_{it}$, they decide whether or not to become informed.

4. If they do not get informed ($e_{it} = 0$), their new belief is formed to be equal to the DeGroot average ($p_{it} = \hat{p}_{it}$), and next period they repeat from 1. If they get informed ($e_{it} = 1$), they now know their true probability distribution (p_i^*) and makes adoption decision based on that. The true probability distribution also becomes their belief from the next period onward ($p_{is} = p_i^*\ \forall s \geq t$).

2.2 Implications of the Model

Consider the situation when there are only two states of the world: one where the new technology has a higher payoff than the traditional one (denoted ω_H), and the other where the new technology has a lower payoff than the traditional one (denoted ω_L). Thus $\Omega = \{\omega_H, \omega_L\}$. Let $p^*_{iH} := p^*_i(\omega_H)$ denote the true probability that for household i the new technology has a higher payoff than the traditional one. Suppose $p^H_{it} := p_{it}(\omega_H)$ is household i's belief of p^*_{iH} at period t. Then, following (1), p^H_{it} is equal to p^*_{iH} if the household is informed, otherwise it is equal to $\hat{p}^H_{it} := \hat{p}_{it}(\omega_H)$. Here $\hat{p}^H_{it}$ denotes the households' approximation of p^*_{iH} based on their network, following DeGroot averaging. Under this simplified scenario, I can now solve the model following backward induction.

In step 2, conditional on being informed, the household decides whether or not to adopt the new technology. The household will adopt the new technology if and only if:

$$\begin{aligned} & p^*_{iH}\pi^N(\omega_H) + (1 - p^*_{iH})\pi^N(\omega_L) - c_i \geq \pi^T \\ & \Rightarrow p^*_{iH} \geq \frac{c_i + (\pi^T - \pi^N(\omega_L))}{(\pi^N(\omega_H) - \pi^N(\omega_L))} =: \bar{p}^*_{iH}. \end{aligned} \tag{4}$$

That is, if and only if the true probability of success with the new technology (p^*_{iH}) is higher than the threshold ($\bar{p}^*_{iH}$), it is profitable for the household to adopt the new technology. The threshold has the cost of switching to the new technology in the numerator and the net benefit of success (compared to failure) with the same technology at the denominator. The cost of switching to the new technology is the sum of direct cost (c_i) and the opportunity cost of switching to the technology only to realize a lower payoff than the traditional technology ($\pi^T - \pi^N(\omega_L)$). Thus, if and only if the probability of success with the new technology is higher than the cost of switching as a fraction of associated benefits, it is optimal for the household to adopt the technology. Given this condition for adoption in step 2, in step 1 the household i will choose to get informed at time t if and only if:

$$\begin{aligned} & p^H_{it}\pi^N(\omega_H) + (1 - p^H_{it})\pi^N(\omega_L) - c_i - \pi^T \geq \eta_i \\ & \Rightarrow p^H_{it} \geq \frac{c_i + (\pi^T - \pi^N(\omega_L))}{(\pi^N(\omega_H) - \pi^N(\omega_L))} + \frac{\eta_i}{(\pi^N(\omega_H) - \pi^N(\omega_L))} =: \bar{p}^*_{iH} + \bar{\eta}_i. \end{aligned} \tag{5}$$

The condition (5) takes the cost of effort (η_i) into account. This is because the decision in step 1 is regarding whether or not to get informed. From (4) and (5), it is clear that if, for household i, p^H_{it} is equal to p^*_{iH}, and they choose to get informed in step 1, they will also adopt the technology in step 2. Conversely, if (4) is not satisfied, then (5) is not satisfied if the diffusion of information is efficient. In other words, under fully efficient information diffusion, only those who would adopt the technology in step 2 would get informed in step 1. Thus, for these households, the following condition must be true:

$$p^*_{iH} \geq \bar{p}^*_{iH} + \bar{\eta}_i. \tag{6}$$

Equation (6) implies that for households that end up adopting the technology, it must be so that their true probability of success justifies the cost of seeking information ($\bar{\eta}_i$) on top of their threshold probability of adoption ($\bar{p}^*_{iH}$). Suppose for household j, that $\bar{p}^*_{jH} + \bar{\eta}_j \geq p^*_{jH} \geq \bar{p}^*_{jH}$. Then even if p^H_{jt} is equal to p^*_{jH}, household j will end up not getting informed about the technology. Hence they will not adopt the technology, even if it is profitable for them to do so. This is due to the positive cost of learning (η_j). This feature is similar to the models of Chandrasekhar et al. (2018) and Banerjee et al. (2018), where the social stigma of information seeking can stop people from learning.

From the above discussion, it is clear that there are multiple possible equilibria for this model. In particular, the equilibrium depends on the households' initial beliefs. If everyone except household i is informed, DeGroot averaging in this set-up will help household i to make the right decision regarding seeking information. The problem, however, arises when most households are uninformed. Of particular interest is the situation when $p^H_{it} \approx 0 \; \forall it$. This situation occurs when everyone believes with certainty that, for them, the new technology yields a lower payoff than the traditional one. In such a scenario, nobody will adopt the new technology even though it may be efficient for some to do so.

Network-based targeting can help in such a scenario. We can target some households (seeds) to improve adoption. The targeted household i will get exogenously informed about their p^*_{iH} at period t. This will get household j to update their $\hat{p}^H_{jt+1}$ if j puts positive weight on i's opinion. Subsequently, this will lead household k to update their $\hat{p}^H_{kt+2}$ if k puts positive weight on j's opinion, and so on. The outcome of this intervention in terms of technology adoption, a few periods down the line, will depend on the initial targeting strategy. In other words, following the initial seeding strategy, the outcomes will vary depending on the path of information diffusion. In such a scenario, for any given targeting strategy, simulations help in attaining the outcomes. These outcomes will then help in understanding the relative effectiveness of different targeting strategies.

In the next section, I measure the relative performance of two types of such targeting strategies using simulations. In doing so, I consider the networks of households that face the decision problem described in this section. I focus on the scenarios where initially $p^H_{it} \approx 0 \; \forall it$, and thus the need for targeting. My simulations provide testable implications that I take to the data in the subsequent sections.

3 Simulations

In this section, I consider households that vary in terms of their *true* probability distributions associated with a new technology. However, connected households are comparable in observable and unobservable characteristics determining these probability distributions. This similarity helps them seek information from their connections. Under such a scenario, I first demonstrate the potential problem for a centrality-based targeting strategy with the example of a specific network. Then I simulate 200 networks to analyze whether the problem persists on average and compare the centrality-based targeting with a probability-based targeting strategy (defined below). I focus on the scenario where the probability distributions differ across households but remain correlated according to the existing social ties. I observe that the relative performance of targeting strategies depends on the level of heterogeneity in the population.

3.1 An Illustrative Example

I start with the example of a specific network of 10 households. The households are heterogeneous concerning their true probability of success associated with a new technology (represented by the p^*_{iH}s).[15] These probabilities matter for the households as the states of the world are drawn independently every period. The probabilities are correlated according to the existing network structure. This correlation introduces the possibility of learning from the network. The distribution of the true probabilities of success is in Figure 2.

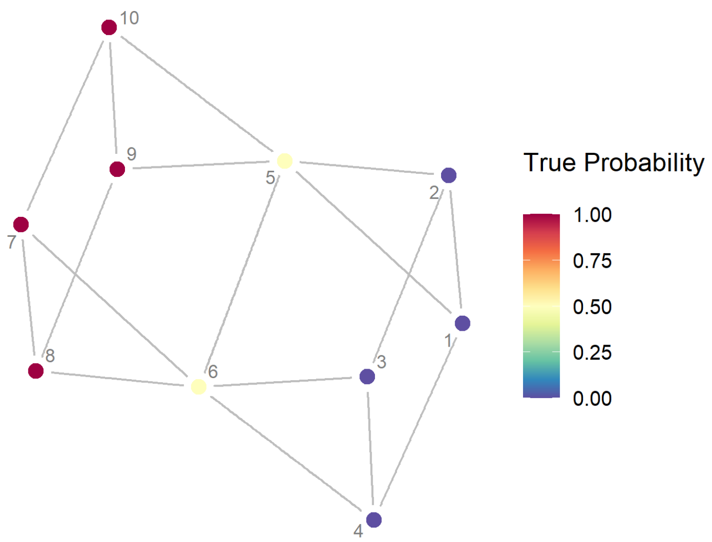


Figure 2: Distribution of True Probability within the network

[15]Here, similar to section 2.2, I am assuming two states of the world: either success or failure with the new technology.

As we can see, there are three types of households in this network. Households with a high probability of success (numbered 7-10), a low probability of success (numbered 1-4), and a moderate probability of success (numbered 5 and 6). For the purpose of this example, consider the threshold probability of learning (i.e. $(\bar{p}^*_{iH} + \bar{\eta}_i)$ in (5)) to be 0.25 for every household. Thus, if the true success probability of a household is more than 25%, the household should get informed under full efficiency. Given the distribution of true probabilities of success shown in Figure 2, it turns out that it is efficient for 6 out of 10 households to get informed in this network (numbered 5-10).

Consider the scenario where, before any interventions, everyone believes their probabilities of success with the new technology to be zero ($p^H_{it} = 0\ \forall it$). Under such a scenario, households will not adopt the technology even if it is optimal for some of them to do so. An intervention is then required to improve adoption. The objective behind such an intervention is to ensure that the households that should adopt the technology with perfect information, end up adopting it. The efficiency of a targeting strategy κ can be measured as:

$$Efficiency_\kappa = \frac{\Psi_\kappa}{\Psi}. \tag{7}$$

Here Ψ denotes the number of informed households under full efficiency. This is the number of households that should get informed as they would adopt the technology under perfect information (i.e. they satisfy equation (6)). For the network in Figure 2, $\Psi = 6$. On the other hand, Ψ_κ captures the number of households that end up getting informed within some periods of implementing the targeting strategy κ, among those that should get informed. For the network in Figure 2, if 2 out of 10 households end up getting informed following the targeting strategy κ, and both of them should have gotten informed under full efficiency, then $\Psi_\kappa = 2$. However, if only one of them should have gotten informed under full efficiency, then $\Psi_\kappa = 1$.

For my analysis, I focus on two types of targeting strategies: centrality-based and probability-based. Similar to BBMM, I seed only two households per network. I consider a centrality-based targeting strategy as the existing literature documents support in favor (Banerjee et al., 2013), and because BBMM recommends centrality-based targeting for the type of diffusion process described here.[16] I consider probability-based targeting as an alternative to centrality-based targeting. The probability-based targeting strategy is to seed households who have the highest true expected benefits with the new technology

[16]The diffusion process described in this paper falls under the category of *complex diffusion*. Complex diffusion models assume that information diffuses to an agent if and only if a certain threshold of the agent's connections gets informed. A more detailed description of different models of diffusion and their use in Development and Agricultural Economics literature can be found in Breza et al. (2019) and Barnett-Howell and Mobarak (2021).

(i.e., the highest p^*_{iH}s in the network). These households are more likely to adopt a technology given a homogeneous cost of learning for everyone and hence are considered to be the early adopters here (definition of early adopters similar to Catalini and Tucker, 2017). The rationale for considering probability-based targeting as an alternative to centrality-based targeting is twofold. First, it is the extreme opposite of the centrality-based targeting strategy. Whereas the centrality-based strategy relies on households that are similar to the average for diffusion, the probability-based strategy does the opposite by focusing on the households more likely to adopt a technology than average. Second, there is a debate in the existing literature regarding whether opinion leaders should be somewhat *superior* to their followers for effective diffusion of new knowledge (Feder and Savastano (2006); Miller and Mobarak (2015)). Through the lens of this debate, probability-based targeting seems to be a natural alternative to centrality-based targeting.

The centrality-based targeting strategy is to seed households central to the network. For the particular example here and the subsequent analysis in this section, I consider centrality in terms of the *eigenvector centrality* measure. The results of my analysis are robust to a different measure of centrality (consult Appendix D for detailed results). Eigenvector-based centrality measures take into account not only the connectivity of a household to other households but also the importance of their connections in terms of their respective connections. Formal definition of different centrality measures can be found in Appendix A.[17] I use the eigenvector centrality measure for two reasons. First, there is evidence in the existing literature in favor of targeting using eigenvector-based measures of centrality (e.g., Banerjee et al. (2013); Beaman et al. (2021a)). Second, for my empirical analysis, I use eigenvector centrality as the primary measure of centrality.

Figure 3 captures the initial seeding for the network from Figure 2 when everyone believes their probabilities of success with the new technology to be zero, thus the need for network-based targeting. In Panel A, seeds are selected based on centrality. Here I seed households numbered 5 and 6, i.e., I consider these households to be informed about their true probabilities of success with the new technology in the first period of policy intervention. Thus, the policy intervention exogenously makes the seeded households aware of their average benefits with the technology. Households 5 and 6 are selected as the seeds because they are the most central households in this network. We can verify that these households are the central-most in this network by counting the number of connections per node. Households 5 and 6 are each connected to five households, whereas every other household in this network has three connections each. Additionally, we can observe that both households have a moderate probability of success with the new technology. This feature is not surprising given that central

[17] For a more detailed description of network centrality measures, consult section 2.2.4 of Jackson (2010) and Bloch et al. (2021).

households are the most connected in the network, and the network is highly assortative in the probability distributions associated with the new technology. Thus, the central households represent the average households in the network, not the ones with a high probability of success with the new technology. This feature has consequences for the final performance of this targeting strategy.

Panel B of Figure 3 capture seeding with probability-based targeting. The seeded households are the ones numbered 8 and 9. These households are selected as they have the highest true probabilities of success with the new technology among all the households in this network. I can easily select these households in simulations, as I can perfectly observe the households' true probabilities of success here. In practice, however, we may not have the information we need to identify these households. In section 4.4, I discuss my strategy for estimating households' true probabilities of success for my empirical analysis. The households selected following a probability-based targeting are, by definition, representing the early adopters in the network. Thus, these households may not be well-connected in the network. This feature has consequences for the final performance of the probability-based targeting strategy.

Panel A: Centrality-Based

Panel B: Probability-Based

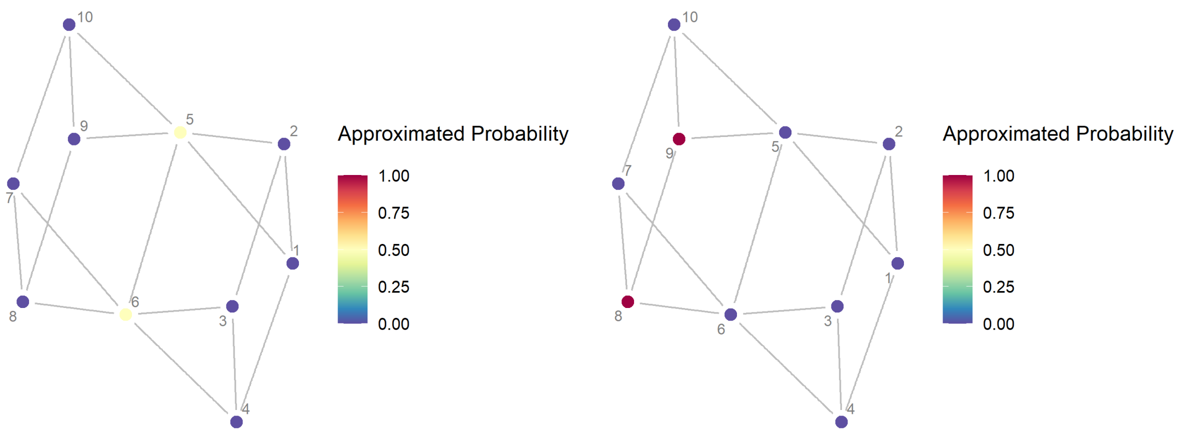


Figure 3: Initial Seeding based on Centrality and Probability

After the initial seeding, I let the diffusion occur over three periods, according to the diffusion process described in the last section. The performance of both targeting strategies at the end of the three periods is in Figure 4. In this particular scenario, probability-based seeds perform better than their centrality-based counterparts. Centrality-based seeds managed to convince no additional households to get informed. On the other hand, probability-based seeds managed to convince everyone else that satisfies equation (6) for this network, to get informed. Using the efficiency measure defined in equation (7), I can score centrality-based targeting 0.33, while probability-based targeting scores 1. Therefore, the centrality-based targeting strategy fails in this scenario.

In this particular example, the probability distributions are highly heterogeneous, and the network connections are highly assortative in terms of these probabilities. The assortativity is necessary for the network to be informative with varying p^*_{iH}s. In what follows, I vary the degree of heterogeneity in p^*_{iH}s, keeping the networks assortative in terms of these probabilities. In particular, for 200 such assortative networks, I vary the correlation level of the p^*_{iH}s. [18] The objective is to note the relative performances of centrality and probability-based targeting strategies over varying degrees of population-level heterogeneity.

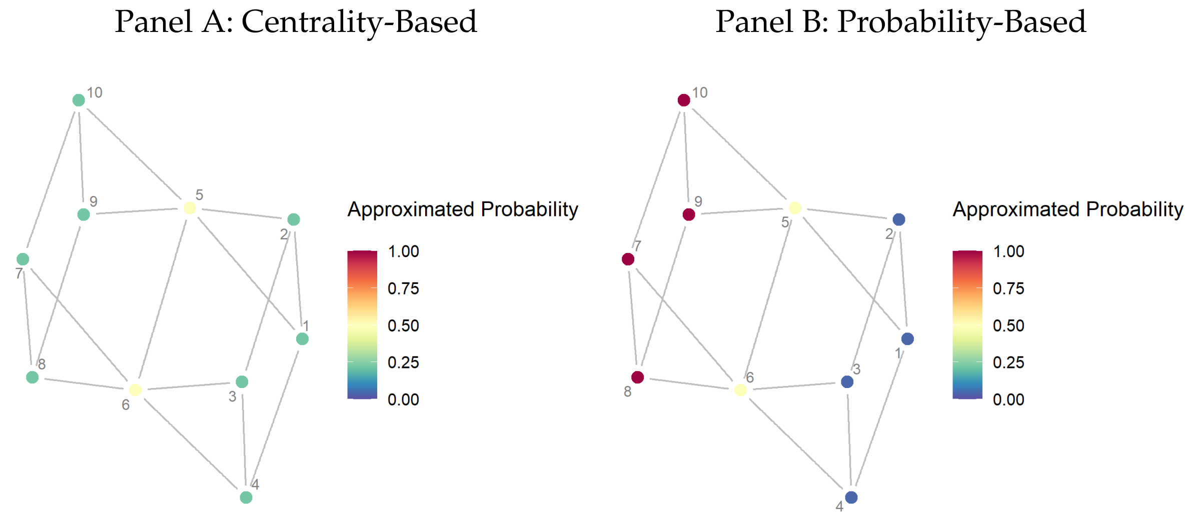


Figure 4: Performance of seeds after three periods

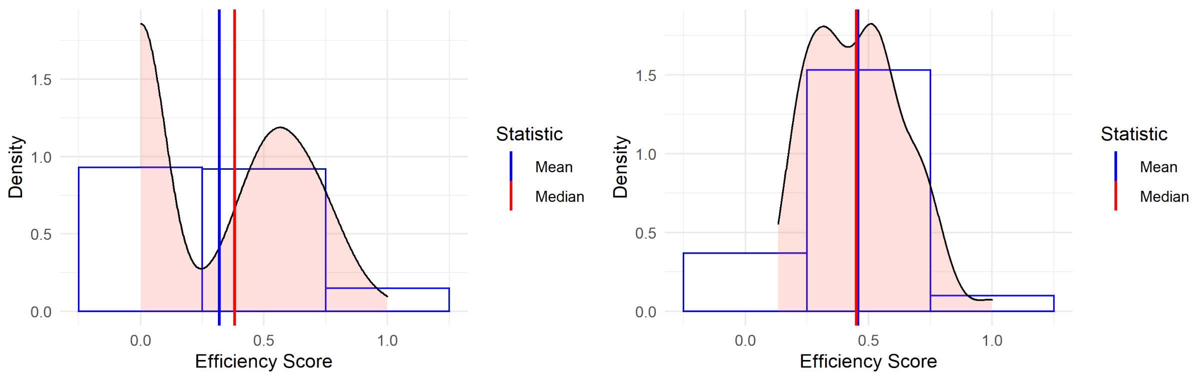
3.2 Targeting Networks with Varying Heterogeneity

What if a new technology is more beneficial to some households than others, with this heterogeneity in benefits affecting the diffusion of information? In this subsection, I explore the performances of different targeting strategies in such a scenario. In particular, I vary the degree of heterogeneity within a network in terms of p^*_{iH}s, and observe the performances of centrality-based and probability-based targeting strategies. I focus on the distribution of efficiency scores after ten periods of simulations over 200 such networks. Table 1 presents the results for three different levels of correlation of p^*_{iH}s between households within networks. Figure 5 represents the corresponding distributions.

The results suggest that if there is a low correlation between the p^*_{iH}s, i.e., the network heterogeneity is high in terms of the benefits from the new technology, probability-based targeting outperforms eigenvector centrality-based targeting.

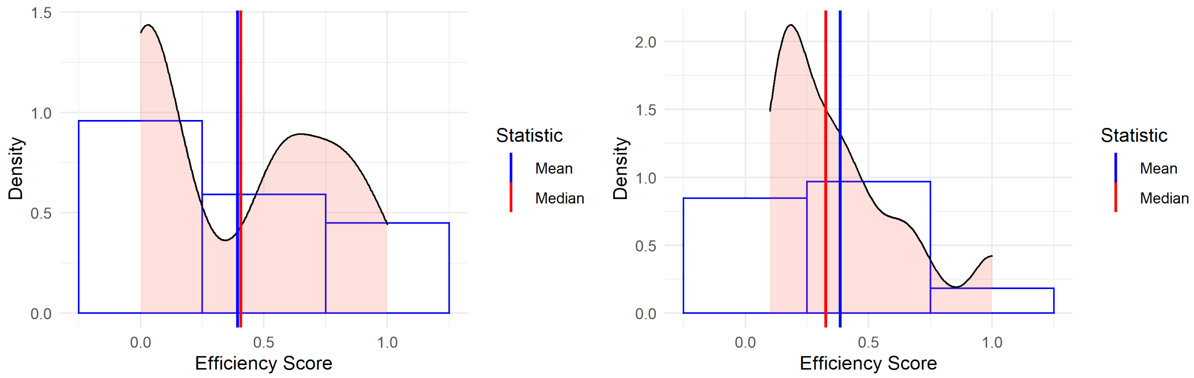
[18] So, in the limit, if the p^*_{iH}s are the same across households, everyone in this network would be connected following the assortative property. In other words, all households would be equally central, and targeting won't be needed. However, the focus is on the cases where the p^*_{iH}s are *sufficiently* different.

Centrality-Based Probability-Based



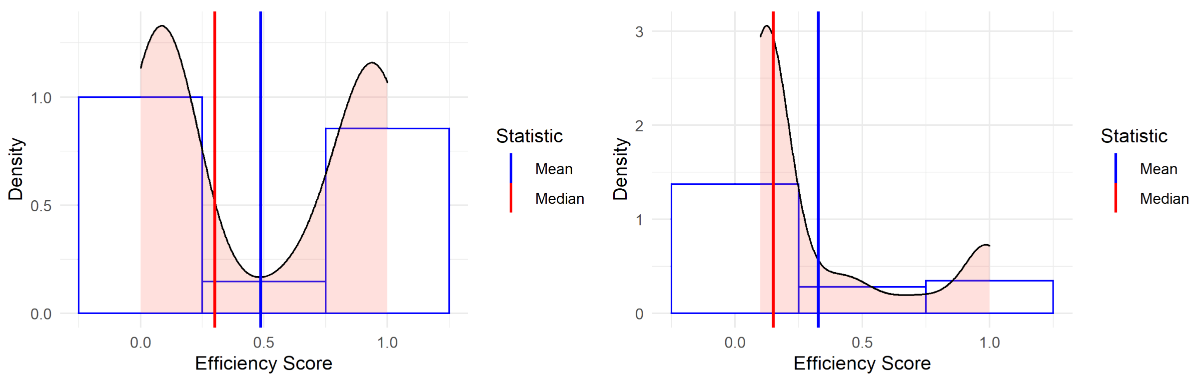
Panel A: Low Level of Correlation

Centrality-Based Probability-Based



Panel B: Medium Level of Correlation

Centrality-Based Probability-Based



Panel C: High Level of Correlation

Figure 5: Distribution of efficiency scores when probabilities are imperfectly correlated

In particular, I observe a higher mean (0.46 vs. 0.32) for probability-based targeting compared to centrality-based targeting. As the level of correlation increases to a medium level, I notice that the gaps between the efficiency of the two types of targeting strategies are closing. More specifically, for a medium level of correlation between the p^*_{iH}s, centrality-based targeting has a mean of 0.39, compared to 0.38 for probability-based targeting. Finally, I consider the results with a high level of correlation between the p^*_{iH}s, which represents a low heterogeneity of the networks in terms of the applicability of the new technology. Here, centrality-based targeting outperforms probability-based targeting in terms of mean efficiency (0.49 vs. 0.33). All of these results are also in Table 1.

Table 1: Simulation Results

		Level of Correlation		
Strategy	Statistic	Low	Medium	High
Eigenvector Centrality-Based	Mean	0.32	0.39	0.49
	Variance	0.10	0.13	0.17
Probability-Based	Mean	0.46	0.38	0.33
	Variance	0.03	0.07	0.11
	Observations	200	200	178

Notes: Simulations on varying levels of correlation are all done for 200 networks, each containing 20 households. However, upon generation of the true probabilities, some networks are dropped as they contained 0% of informed households under full efficiency.

3.3 Discussion

From the above analysis, I observe that the relative performance of centrality-based and probability-based targeting strategies in diffusing knowledge regarding a new technology depends on the heterogeneity within a network. In particular, the heterogeneity of interest is in the benefits of the technology. When the heterogeneity is high, there is a clear distinction between the households that should adopt the technology and the ones that should not. In such a scenario, targeting households that are central to the network is not efficient. Targeting households more representative of the households that should adopt the technology works better if the network is highly assortative in terms of characteristics that determine the benefits. On the contrary, if everybody in the network is similar in the benefits of adoption, centrality-based targeting works better. This is because, for a homogeneous population, central seeds maximize diffusion for a complex diffusion process, as documented by BBMM.

Figure 6 represents the mean efficiency scores over a range of different levels of correlation in p^*_{iH}s. The main result discussed above holds for this wider range of observations. As the correlation increases, heterogeneity in p^*_{iH}s decreases, centrality-based targeting performs better, and probability-based targeting performs worse. This result is robust to a different measure of centrality (discussed in Appendix D).

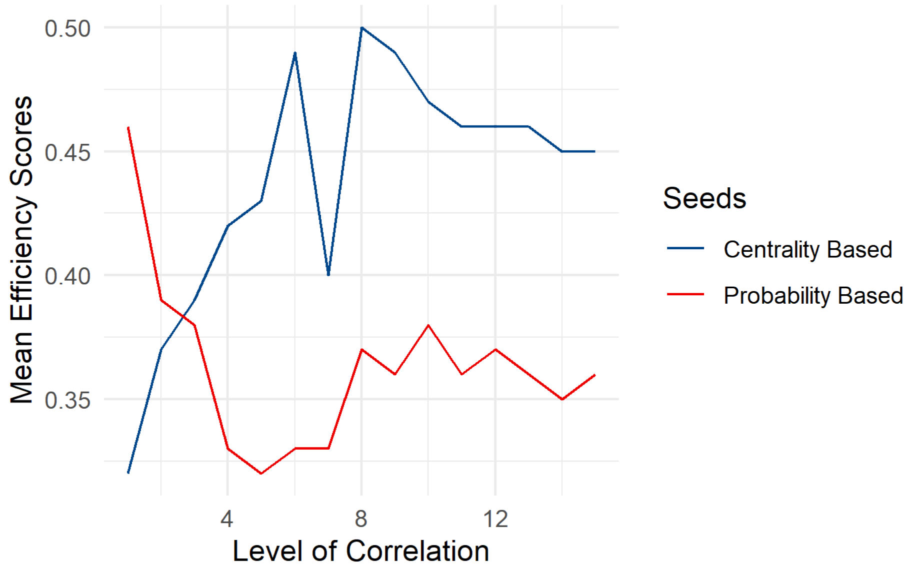


Figure 6: Mean efficiency scores over increasing levels of correlation

A higher mean efficiency score indicates better performance on average. Thus, the centrality-based strategy is *better* for high levels of correlation; the probability-based strategy is *better* for low levels of correlation. These results lead to the main hypotheses of my study that I take to the data in the subsequent sections. For my empirical strategy, it helps to think of the two targeting strategies with respect to a benchmark. In particular, consider the benchmark of selecting two seeds randomly. I expect the efficiency scores of a random assignment to remain the same over different heterogeneity levels due to random assignments being independent of the correlation in p^*_{iH}s. Thus, with respect to this benchmark, I expect probability-based targeting to perform better and centrality-based targeting to perform worse, as heterogeneity in p^*_{iH}s increases.

4 Empirical Framework

My next objective is to empirically test the following set of hypotheses derived from my theoretical framework using simulations.

Hypotheses: As the level of heterogeneity in terms of the benefits of a new technology increases:

1. central seeds perform worse in terms of diffusing that technology.

2. probability-based seeds perform better in terms of diffusing that technology.

It is important to note that Hypothesis 1 does not require networks to be highly assortative in terms of characteristics determining the heterogeneity in benefits. Even if the networks are not assortative, as long as the heterogeneity affects the diffusion process, I expect the hypothesis to be true. Thus, if Hypothesis 1 is true, it will shed light on the underlying condition for the failure of centrality-based targeting even with a complex diffusion process. Hypothesis 2, on the other hand, require assortativity in the network in terms of characteristics determining the heterogeneity in benefits. Thus, accepting Hypothesis 2 would mean the existence of such assortativity in the networks that can be useful for policy purposes.

4.1 The Ideal Experiment

Before diving into the data and the description of identification strategies used in this study to validate my theoretical findings, it is helpful to think about the ideal experimental setup for the validation. In an experiment, I need to randomly allocate regions/villages into three types of seeding strategies:

1. Centrality-based seeding

2. Probability-based seeding

3. Random seeding (which will serve as a benchmark/ control group)

Then I can run the following reduced-form regression:

$$\begin{aligned} Y_v = \alpha_0 + \alpha_1 \textit{Centrality Based}_v + \alpha_2 \textit{Probability Based}_v + \alpha_3 \textit{Heterogeneity}_v \\ + \alpha_4 \textit{Centrality Based}_v \times \textit{Heterogeneity}_v + \alpha_5 \textit{Probability Based}_v \times \textit{Heterogeneity}_v + u_v. \end{aligned} \tag{8}$$

Here Y_v denotes diffusion related outcome variable of interest for village v. The dummies *Centrality Based*$_v$ and *Probability Based*$_v$ indicate whether the village got assigned to either centrality or probability-based seeding strategy. *Heterogeneity*$_v$ is the village-level Coefficient of Variation (CV) for the probability of adoption, capturing the village-level heterogeneity in terms of the benefits of the new technology. Finally, u_v is a random error term in the regression. My hypotheses state that $\alpha_4 < 0$ and $\alpha_5 > 0$.

4.2 Data Sources

I do not have access to the data from the ideal experiment described above. Hence, I use the replication data from BBMM together with the survey data from AESTAS conducted by IFPRI. In this subsection, I describe these datasets before proceeding to the description of my identification strategies in the next subsection.

4.2.1 Replication data of BBMM

BBMM conducted a Randomized Controlled Trial (RCT) to promote *Pit Planting* (PP) for Maize farmers in Malawi.[19] The researchers seeded 200 villages from 3 Malawian districts with semi-arid climates (Machinga, Mwanza, and Nkhotakota) with 2 'seed' farmers each. The objective was to induce widespread social learning of PP. The intervention involved training the seed farmers on PP (and CRM), with the material of training remaining the same across different treatment arms. The villages were equally divided into four experimental groups:

1. **Complex Contagion:** Seeding done assuming the underlying diffusion process to be of complex diffusion. Under the assumption of this diffusion process, the information diffuses only if a certain threshold of each household's connections gets informed. Under this assumption, both the optimally chosen seeds were central in the network.

2. **Simple Contagion:** Seeding done assuming the underlying diffusion process to be of simple diffusion. Under the assumption of this diffusion process, the information diffuses with a random probability from one household to its connections. Under this assumption, the optimal choice was to have one central seed household and one seed household on the periphery.[20]

3. **Geo:** Villages were seeded solely based on geographic proximity. As a result, the seed households were geographically located near each other but were not central (in the network data).

4. **Benchmark (control):** Extension agents selected two seeds like they usually do.

It is important to note that this experimental set-up focuses on seeding households solely based on their positions in the network (in terms of social ties or geographic ties). Thus, the diffusion of information was assumed to be independent of other

[19] They also promoted *Crop Residue Management* (CRM). However, the sample on the use of CRM is small. Thus, similar to the main analysis of BBMM, I focus on PP only. I also do not expect my predictions to be valid for CRM, as CRM is not a *new* technology in the sampled areas. However, PP is a fairly new technology there, so I expect my predictions to hold for PP.

[20] Households on the periphery of a network represent households that are not well connected in terms of existing social ties.

household characteristics. On the contrary, I consider households to be heterogeneous in their expected benefits from the new technology, with this heterogeneity affecting the diffusion of information for a given seeding strategy.

The researchers first collected the social network census data in 2010-11, before any intervention or household survey took place. The census elicited names of people each respondent consults when making agricultural decisions, information on household composition, socio-economic characteristics of the household, general agriculture information, and workgroup membership information. They matched these responses with the village listing to identify links. They considered individuals linked if either party named each other (undirected network) or if they are part of the same household. Based on this network information, the researchers used simulations to identify seeds according to the different diffusion processes to optimize diffusion after four periods. For each of the 200 villages in their study, the researchers used the simulations to identify the optimal choice of 2 seeds following complex diffusion, simple diffusion, and geographic proximity. The villages were then randomly allocated to one of the four treatment groups. Depending on the allocation, 2 seed households were selected per village. The researchers asked extension agents to identify benchmark seeds only for the villages allocated to the control group. The seed households then received training on PP (and CRM). Once the training was complete, the researchers conducted household surveys to collect data on farming techniques, input use, yields, assets, and other characteristics.

The researchers randomly surveyed a panel of approximately 30 households per village, involving all the seed and shadow farmers, along with 22-24 other farmers.[21] They collected information on approximately 5600 households from the 200 villages. In 2 districts (Machinga and Mwanza) that consist of 141 study villages, they collected three rounds of survey data in 2011, 2012, and 2013. Due to unanticipated delays in project funding, in the third district (Nkhotakota), they could only start the operation in 2012. Hence, for the third district with 59 study villages, they collected only two rounds of survey data (in 2012 and 2013). The first round of the survey was conducted a few months after the training of the seed farmers. This round attempted to capture some baseline characteristics and knowledge levels of the surveyed households regarding PP (and CRM). Every survey round was conducted at the start of the agricultural season, after the land preparation. As PP is used for land preparation, the households' adoption decision of PP was observed three times for Machinga and Mwanza, and twice for Nkhotakota.[22] For more details on the intervention and sampling of the study, please consult BBMM.

[21]Shadow farmers are seed farmers chosen by the simulation, assuming some underlying diffusion model. But they were not selected as seeds due to their villages getting assigned to receive seeding based on a different diffusion model.

[22]Similarly, since CRM is used after harvest, the adoption of CRM was observed only twice for Machinga and Mwanza, and once for Nkhotakota. Thus, the sample on the use of CRM is limited.

The objective of BBMM is to assess the effectiveness of different centrality-based targeting strategies on the adoption of pit planting. For that purpose, they collected detailed data on household-level adoption decisions over multiple survey rounds. The replication package also includes information on household-level measures of centrality used to select seeds under different experimental interventions. The former helps me calculate the dependent variables for my analysis, while the latter helps by providing the information I require to assess the centrality of seed households in the experiment. Additionally, I need the surveyed households' ex-ante probability of adoption for my analysis. This information is not available in the replication data as BBMM does not consider the benefits of adoption to be different across households. For this purpose, I turn to the AESTAS dataset.

4.2.2 AESTAS data

AESTAS is a nationally representative household survey conducted by International Food Policy Research Institute (IFPRI). The objective of this survey was to monitor the lead farmer (LF) program in Malawi.[23] The survey covered all 29 districts of Malawi, except Likoma.[24] The data collection was done in two waves: wave 1 in 2016 and wave 2 in 2018. The publicly available version of the survey dataset contains information from three different types of interviews:

1. **Household Interviews:** Ten households were randomly chosen for interview from randomly selected sections within each district.[25] Stratification was done based on whether or not the household had a LF. Per section, up to two households with LFs were selected. A total of around 299 sections were surveyed. The same households were interviewed in the two waves with a small level of attrition (around 4%). Around 3000 households were surveyed in wave 1, with 2880 among them being re-surveyed in wave 2. For each household, both the household head and their spouses were interviewed. The survey collected data on technology adoption and awareness, exposure to different technologies, access to extension services, and socioeconomic and demographic characteristics.

2. **Lead Farmer (LF) Interviews:** Around 531 LF households were selected for household interviews. During the first wave of the household survey, these LF households were asked additional questions. These questions collected information on the LF's characteristics, activities, roles, expectations, incentives, challenges, suggestions, any support they receive from agricultural extension development officers (AEDOs) and other organizations, etc.

[23]Consult Khaila et al. (2015) for details on the lead farmer program.

[24]The survey considered the Mzimba district as divided into North and South, and the Lilongwe district as divided into East and West.

[25]Sections are geographical units in Malawi that are one level lower than districts.

3. **Community Interviews:** In addition to the household surveys, 2-4 leaders per village were surveyed in both waves. The objective was to collect community-level information like the number of lead farmers, type of training they received, number of projects, and other community characteristics.

More information on the survey and associated sampling can be found in Ragasa and Niu (2017), Niu and Ragasa (2018), Ragasa (2020), and Ragasa et al. (2021).

For this study, I use the data collected through household interviews only. In particular, I am interested in the data on household-level technology adoption. Two types of technology adoption information are available in the data:

1. Reported adoption for a list of pre-determined technologies and practices. This list focuses on both agricultural and food processing practices.

2. Reported plot-level usage for a list of pre-determined agricultural technologies and practices.

This information helps me calculate adoption indices crucial to my analysis (see Appendix C for details on the construction of these indices). I use these indices as proxies for the probability of adoption.

4.3 Identification Strategies

I now turn to discuss the identification strategies of my empirical analysis. The identification uses within and between treatment group variations in the BBMM sample. The objective is to use the experimental setup of BBMM to test the predictions of my simulations. Contrary to BBMM's focus on comparing the effectiveness of different centrality-based targeting strategies, I focus on assessing the efficacy of centrality-based targeting vis-à-vis probability-based targeting for varying degrees of population heterogeneity.

In this subsection, I first discuss how I explore the overall village-level variations in the data. These are non-experimental variations. Thus, identification using them requires some assumptions. I discuss these assumptions in detail. Next, I focus on the identification using experimental variations. Both identification strategies require the calculation of the probability of adoption at the household level. For that purpose, I use the survey data from AESTAS. The last subsection of this section provides details on that.

4.3.1 Exploring village level variations

Given the selection of seeds in the BBMM experiment, I calculate the seeds' average centrality and probability of adoption. This information is used in the following regression:

$$
\begin{aligned}
Outcome_{vt} = {} & \beta_0 + \beta_1 Centrality_v + \beta_2 Probability_v + \beta_3 Heterogeneity_v \\
& + \beta_4 Centrality_v \times Heterogeneity_v + \beta_5 Probability_v \times Heterogeneity_v + \lambda X_v + \zeta_t + \epsilon_{vt}.
\end{aligned} \tag{9}
$$

$Outcome_{vt}$ denotes some adoption-related outcome for village v at time t. Like BBMM, I focus on the outcomes in years 2 and 3. I discuss the outcome variables of my analysis in the next section. $Centrality_v$ represents the average centrality of the seeds for village v, at the baseline. I calculate this by using the eigenvector centrality of the seed households at the baseline. The centrality measures are pre-calculated by BBMM and available in their replication data. $Probability_v$ represents the average probability of adoption for the seeds in village v, at the baseline. I proxy for the probability of adoption using predicted adoption and usage indices. I calculate these indices at the baseline, conditional on some observable household demographics. The calculation uses estimates from another regression, which I discuss in the next subsection. I use the coefficient of variation of the same adoption and usage indices at the village level to capture village-level heterogeneity in the benefits of a new technology. This is represented by the $Heterogeneity_v$ variable in (9). It is important to note that both the probability of adoption and the related coefficient of variation are proxied by variables that are calculated conditional on observable demographics. These variables are therefore not particular to any technology. Instead, they represent whether the households are likely to adopt any new technology conditional on their observable characteristics. Following my hypothesis, I expect $\beta_4 < 0$ and $\beta_5 > 0$. I control for baseline village level characteristics (X_v), and year-fixed effects (ζ_t). The random error of the regression is captured by ϵ_{vt}.

The calculation of outcome variables excludes the seeded households. BBMM use the same outcome variables in their village-level analysis. I calculate $Centrality_v$ and $Probability_v$ using the information on the seeded households. I assume that seed household characteristics are exogenous to the outcome variables. The assumption seems reasonable as the village-level outcomes do not consider the seeded households. I assume that, conditional on these village level controls, $Heterogeneity_v$ is also exogenous in (9). In any case, my coefficients of interest are β_4 and β_5. As long as $Centrality_v$ and $Probability_v$ remain exogenous, I do not need to take any stand regarding the exogeneity of $Heterogeneity_v$ for identifying my coefficients of interest.

As defined above, $Centrality_v$ represents the average centrality of seed households at the village level. Endogenous $Centrality_v$ in (9) implies unobserved village-level characteristics correlating with the network positions of the seed households and the village-level outcomes calculated excluding the seed households. For example, there may be unobserved social learning correlating with the network positions of the seeds and the adoption-related outcomes. However, this is more likely to be true for the household-level outcomes. At the village level, unless there is a village-level learning process correlating with the seed households' network positions, $Centrality_v$ should be exogenous in (9). Similarly, the village-level unobserved characteristics affecting adoption-related outcomes should not be related to the seed's adoption probability. As $Probability_v$ represents the average adoption probability of the seed households, it should also be exogenous in (9).[26] However, since I cannot verify these identifying assumptions, I also use the experimental variations in the BBMM data that use weaker identifying assumptions. The next subsection provides details on that.

Finally, not accounting for the treatment status in the regression can lead to omitted variable bias if there is some measurement error in calculating $Centrality_v$, as the experimental design ensures that some villages will have more central seeds than the other. In Appendix D, I check the robustness of my results by including the treatment dummies. As my results remain almost the same, in the following section I present them without the treatment dummies.

4.3.2 Exploring between treatment group variations

To explore between treatment group variations, I use the following specification:

$$\begin{aligned}
Outcome_{vt} &= \theta_0 + \theta_1 Centrality_v + \theta_2 Probability_v + \theta_3 Heterogeneity_v \qquad (10)\\
&+ \xi_b Centrality_v \times Heterogeneity_v + \xi_c Centrality_v \times Heterogeneity_v \times Complex_v\\
&+ \xi_s Centrality_v \times Heterogeneity_v \times Simple_v + \xi_g Centrality_v \times Heterogeneity_v \times Geo_v\\
&+ \phi_b Probability_v \times Heterogeneity_v + \phi_c Probability_v \times Heterogeneity_v \times Complex_v\\
&+ \phi_s Probability_v \times Heterogeneity_v \times Simple_v + \phi_g Probability_v \times Heterogeneity_v \times Geo_v\\
&+ \gamma X_v + \rho_t + \eta_{vt}.
\end{aligned}$$

Specification (10) is similar to specification (9), except the interactions of $Centrality_v \times Heterogeneity_v$ and $Probability_v \times Heterogeneity_v$ with treatment dummies. Here, ξ_b captures the interaction between seed centrality and village level heterogeneity for the benchmark treatment group. ξ_c, ξ_s, and ξ_g captures how that interaction changes compared to the benchmark for complex, simple, and geo treatment groups. Similarly, ϕ_b captures the interaction between seed probability and village level heterogeneity for the benchmark treatment group. ϕ_c, ϕ_s, and ϕ_g captures how that interaction changes

[26] An example of a village-level unobserved learning process correlating with the seed households' network positions (or adoption probability) would be when the seeds with higher centrality (or probability) are more likely to broadcast information to the masses affecting village-level adoption. Not controlling for this information will make $Centrality_v$ (or $Probability_v$) endogenous in (9).

compared to the benchmark for complex, simple, and geo treatment groups. Thus, for example for complex treatment group, the effect of $Centrality_v \times Heterogeneity_v$ on the outcome variable is $(\xi_b + \xi_c)$; the effect of $Probability_v \times Heterogeneity_v$ on the outcome variable is $(\phi_b + \phi_c)$. I expect the impact of $Centrality_v \times Heterogeneity_v$ to be negative and the effect of $Probability_v \times Heterogeneity_v$ to be positive, within different treatment groups. However, using this specification, I am more interested in exploring between group variations. Thus, main coefficients of interest in this specification are: $\xi = \{\xi_c, \xi_s, \xi_g\}$ and $\phi = \{\phi_c, \phi_s, \phi_g\}$.

For a treatment group having the same level of heterogeneity as the benchmark, I expect outcomes to be positively related to centrality and negatively related to probability. Thus, given the population heterogeneity of a group and the adoption probability of the seeds, moving to higher (lower) centrality seeds helps diffuse the technology to more (less) households. Similarly, given the population heterogeneity of a group and the centrality of the seeds, moving to higher (lower) probability seeds diffuse the technology to fewer (more) households. I argue the same for (9) while exploring the village-level non-experimental variations. Thus, I skip the reasoning of this argument here.

If the treatment group is less heterogeneous than the benchmark, I expect seeds with higher centrality to perform better and seeds with a higher adoption probability to perform worse. From my simulations, I expect centrality-based targeting (or probability-based targeting) to perform better (worse) with population homogeneity. If a treatment group is less heterogeneous than the benchmark, it is more homogenous in its population's probability of adoption. Thus, I expect more central seeds to perform better and seeds with higher adoption probability to perform worse. In this case, however, my theory does not have any prediction for seed households with lower centrality and adoption probability. For treatment groups less heterogeneous than the benchmark, the effect of having seeds with less centrality (or less adoption probability) depends on the relative impacts of the population homogeneity and centrality (or adoption probability). Hence, I have no specific predictions on the performance of such seeds. Similarly, for treatment groups having higher population heterogeneity than the benchmark, I expect the seed households with lower centrality to perform better and seeds with lower adoption probability to perform worse. In this case, my theory does not have any prediction for the seed households with more centrality and adoption probability.

As an example, let us consider the complex treatment group. If this group has the same level of heterogeneity as the benchmark, I expect ξ_c to be positive (negative) if the complex treatment group has more (less) central seeds than the benchmark. Similarly, I expect ϕ_c to be negative (positive) if the complex treatment group has a higher (lower) seed adoption probability than the benchmark. Now, if the complex treatment group is less heterogeneous than the benchmark, I expect the following:

- If they have more central seeds than the benchmark: positive ξ_c; less central seeds than benchmark: depends on the relative effects of the drop in centrality and heterogeneity.

- If they have seeds with higher adoption probability than the benchmark: negative ϕ_c; seeds with a lower adoption probability than benchmark: depends on the relative effects of the drop in probability and heterogeneity.

Similarly, if the complex treatment group is more heterogeneous than the benchmark, I expect:

- If they have less central seeds than the benchmark: positive ξ_c; more central seeds than benchmark: depends on the relative effects of the increase in centrality and heterogeneity.

- If they have seeds with a lower adoption probability than the benchmark: negative ϕ_c; seeds with higher adoption probability than benchmark: depends on the relative effects of the increase in probability and heterogeneity.

Like specification (9), I control for baseline village level characteristics and year-fixed effects. As the coefficients of interest use interactions with the treatment dummies, I do not need any additional assumption other than assuming the success of the randomization. Finally, it is important to note that I do not include the treatment dummies in specification (10) as BBMM argues that the treatment status affects the outcome variables only through the centrality of the seeds. In Appendix D, I present the robustness of my results by including the treatment dummies. My results remain robust.

4.4 Strategy for Approximating Probabilities of Adoption

For my regression specifications, I need to calculate the probability of adopting a new technology for all households. The average of this probability measure for seed households is $Probability_v$ in the regressions, while the coefficient of variation of this measure at the village level is $Heterogeneity_v$. However, BBMM did not collect any information about these probabilities, as their micro-foundations assumed benefits from the new technology to be the same across households. Hence, I need to approximate these probabilities conditional on the observable characteristics of the households surveyed in their study.

For this purpose, I use the data from AESTAS. The data contains information on technology adoption and household characteristics. It surveys a nationally representative set of farmers in Malawi on a universe of technologies that includes the technologies covered in BBMM. I use this information on the universe of technologies to calculate adoption and usage indices. Appendix C contains details on the construction of these indices. I use the following regression specification to estimate the mapping from observable household characteristics to the adoption index:

$$Adoption\ Index_{it} = f(X_{it}; \mu_{it}), \tag{11}$$

where X_{it} are observable household characteristics. I consider only the characteristics observed in both AESTAS and BBMM data. I present the robustness of the regression results in the next section to other household characteristics observable only in the AESTAS data and not in the BBMM data. The term μ_{it} captures the random error in the regression. In my preferred specification, I consider function $f(\cdot)$ to be linear (thus, the estimation uses ordinary least square). However, I check the robustness of my results to non-linear specifications. I present these in Appendix D. I use a similar regression specification to estimate the mapping from observable household characteristics to the usage index.

I use the estimations of this model to construct the adoption index (and the usage index) conditional on the observable demographics in the BBMM dataset. I use this variable to proxy for the households' adoption probability. We should note that (11) gets estimated with possible omitted variable bias. For example, there may be possible social learning correlating with both the adoption index and observable demographics.[27] Thus, the coefficients estimated using (11) would represent a correlation, not causality. This bias in estimating households' adoption probabilities should not affect my coefficients of interest in (10), as the identification uses experimental variations. However, we must consider the consequences for (9). The bias in estimating households' adoption probabilities would lead to a biased $Probability_v$ in (9). However, this will only create a problem in identifying the coefficient of interest β_5 if this bias correlates with unobserved village-level characteristics affecting adoption-related outcomes. This correlation is less likely to be true because:

1. Household level bias should not correlate with village-level unobservables.

2. Bias in the estimates originating from the AESTAS sample should not correlate with the unobserved village-level variations in the BBMM sample.

However, since I cannot verify these assumptions, specification (10) provides a better alternative.

[27]More specifically, in the AESTAS data, the households having higher adoption index may adopt more technologies due to being connected to the lead farmers. Not controlling for this regression will over-estimate the adoption index for their demographics.

5 Results and Discussion

In this section, I present the empirical results of my analysis. I start by discussing how I approximate adoption probabilities using the AESTAS data. In doing so, I present the associated regression results and the assumptions I need for using these results for the rest of my analysis. The following subsection explores the main variations in the BBMM data via descriptive statistics. In the final sub-section, I present the main empirical results of this study.

5.1 Approximating Probabilities of Adoption

Table 2: Baseline Demographics Across Datasets

		Variables				
Dataset	Statistic	Adults	Children	Housing	Livestock	Assets
AESTAS	Mean	2.14	3.00	-0.09	-0.03	-0.03
	(SD)	(1.00)	(2.00)	(0.98)	(0.99)	(1.00)
	Median	2.00	3.00	-0.29	-0.40	-0.29
	Observations	2820	2820	2803	2820	2820
BBMM	Mean	2.36	2.77	-0.02	0.02	0.09
	(SD)	(0.95)	(1.86)	(0.99)	(1.02)	(1.03)
	Median	2.00	3.00	-0.24	-0.31	-0.10
	Observations	5384	5407	5382	5407	5407

Notes: The variables *Adults* and *Children* represent number of adults and children in a household, respectively. The variables *Housing*, *Livestock*, and *Assets* were standardized first principal components. For the AESTAS sample: *Housing* includes information on materials walls are made of, roof materials, and floor materials. Each of the three variables are coded to be 0- Traditional, 1- Modern. *Assets* includes the number of bicycles, radios and cell phones the household owns. *Livestock* includes the number of sheep, goats, chickens, cows, and pigs. For the BBMM sample: *Housing* includes information on materials walls are made of, roof materials, floor materials and whether the household has a toilet. *Assets* includes the number of bicycles, radios and cell phones the household owns. *Livestock* is an index including the number of sheep, goats, chickens, cows, pigs, guinea fowl, and doves. (footnote 1 from Table A5 of BBMM)

I start by comparing key baseline demographic information across datasets. This is presented in Table 2. The comparison is important as it helps me understand how the results derived using the AESTAS data map into the BBMM data. The five variables chosen are available in both AESTAS and BBMM data. In terms of the mean and median, both datasets are similar in the number of adults and children in the household.

However, the BBMM sample is slightly richer than its AESTAS counterpart. We can see this by comparing the mean and median of standardized housing, livestock, and assets PCA (Principal Component Analysis) scores. This is not surprising given that AESTAS focused on a nationally representative sample of farmers in Malawi, whereas BBMM focused only on the Maize farmers.

Table 3: OLS Regression Results for Adoption and Usage Indices

	Adoption Index			Usage Index		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Adults	0.008***	0.008***	0.005**	0.011***	0.011***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Children	0.003***	0.003***	0.002	0.003***	0.003***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Housing	0.009***	0.009***	0.008***	0.003	0.003	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Livestock	0.010***	0.011***	0.005*	0.014***	0.014***	0.009***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Assets	0.024***	0.024***	0.017***	0.020***	0.020***	0.014***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Year Fixed-Effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	5610	5608	5604	5610	5608	5604
R-squared	0.096	0.096	0.150	0.085	0.131	0.169

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors clustered at the section level are in parentheses. All regressions use sample weights and include a constant term. The variables *Adults* and *Children* represent number of adults and children in a household, respectively. The variables *Housing*, *Livestock*, and *Assets* were standardized first principal components. *Housing* includes information on materials walls are made of, roof materials, and floor materials. *Assets* includes the number of bicycles, radios and cell phones the household owns. *Livestock* includes the number of sheep, goats, chickens, cows, and pigs. Household Controls include: gender and age of household head, activity of household head (0- Non-Farmer, 1- Farmer), whether the household applied for a loan in the past, the households' time and risk preferences, and whether a household member is a lead farmer (LF).

Table 3 presents the main results for this subsection. Here, I estimate the adoption and usage indices conditional on the demographics presented in Table 2. The estimation uses the AESTAS data. The first three columns present the results for the adoption index. I observe a positive correlation between the households' wealth level and their adoption index.[28] In addition, families with more adults and children report a higher adoption index. The results remain almost identical when I control for the year fixed effects.

[28]Here households' wealth level is captured by their housing, assets, and livestock principal component analysis scores. Details on these variables are in the footnote of the table.

The magnitudes and significance levels vary, controlling for other household-level characteristics. However, the signs remain the same. The remaining three columns of the table present the results for the usage index. The results are qualitatively similar to that of the adoption index. The most notable difference is that the coefficient corresponding to the housing PCA score is statistically insignificant throughout specifications. The main takeaway from these results is that the coefficients remain similar with or without controlling for year fixed-effects and other household-level characteristics. Thus for calculating the predicted adoption and usage indices, I use the estimates reported in columns (1) and (4), respectively.

Figure 7 compares the actual and predicted indices for the AESTAS sample. The estimates capture only a fraction of the actual variation. The actual adoption index has a mean of 0.085 with a standard deviation of 0.120. In comparison, its predicted counterpart has a mean of 0.086 with a standard deviation of 0.038. The numbers are similar for the usage index in terms of prediction quality. The actual usage index has a mean of 0.163 with a standard deviation of 0.122, whereas the predicted usage index has a mean of 0.162 and a standard deviation of 0.035. Thus, the predictions are good at predicting the mean but only capture a third of the actual variation. This is not surprising given that the predictions are made based on only a few observable demographics.

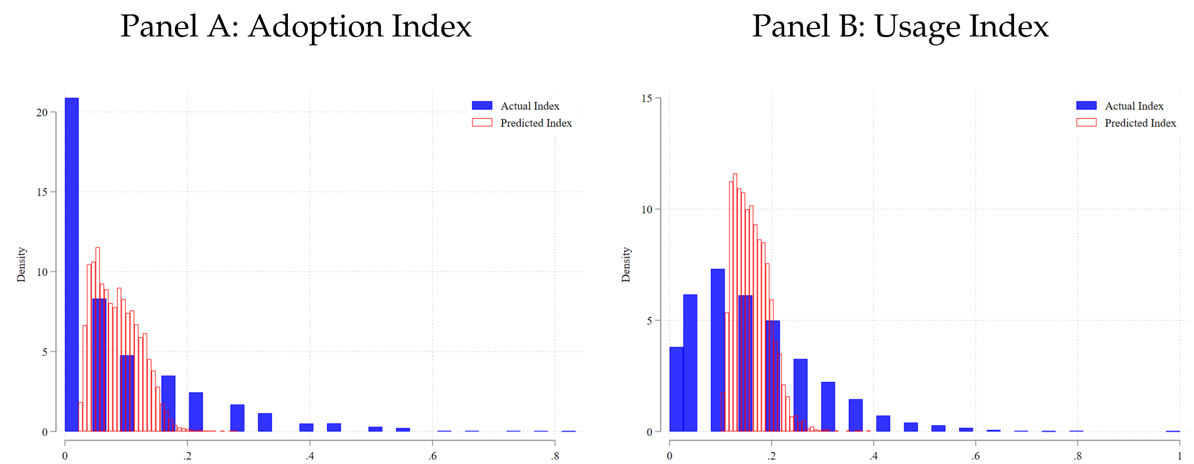


Figure 7: Actual and Predicted Adoption and Usage Indices

To use the estimated mapping from the observable demographics to the adoption and usage indices for predicting the probability of adoption in the BBMM data, I need the following assumptions:

- **Assumption 1:** Adoption and Usage indices are good proxies for the probability of adoption.
- **Assumption 2:** The variation in adoption and usage indices, conditional on the

demographics observable in both AESTAS and BBMM data, is sufficient for my analysis.

- **Assumption 3:** The mapping of observable characteristics to the adoption probability is the same across the datasets I use in this study.

- **Assumption 4:** Any bias in the estimated relationship between adoption probability and observable characteristics is independent of the unobserved village-level learning in the BBMM sample.

The first three assumptions are necessary for extrapolating the AESTAS information to the BBMM data. There is no formal way of testing these assumptions. I need the fourth assumption to identify β_5 in (9), as I already discussed in the last section.

5.2 Descriptive Statistics

Table 4 focuses on describing key baseline characteristics in the BBMM sample. The last column of this table represents overall village-level non-experimental variations. I exploit this variation in the regression specification (9). The first four columns of the table represent the within treatment group variations. Regression specification (10) uses the experimental variations between these four groups. The first two rows present the main outcome variables of my analysis. Adoption Rate (PP) captures the proportion of *typical* farmers per village that adopted pit planting in each agricultural season. Here, *typical* farmers correspond to the farmers that were not selected as *seed* or *shadow* farmers in the experiment. Any Non-Seed Adopters (PP) is a dummy variable that captures whether the villages had at least one *non-seed* farmer adopting pit planting in an agricultural season. I focus on these two outcome variables as BBMM also uses these variables as outcomes in their village-level analysis. The baseline data suggest an adoption rate of around 2-3% across treatment arms. Also, only 30-42% villages had at least one *non-seed* farmer adopting pit planting in the baseline. These numbers suggest low adoption of pit planting in the baseline, providing an ideal setting to test the predictions of my theoretical analysis. Through the lens of my theoretical framework, the pessimism regarding the prospect of pit planting was responsible for the low adoption of pit planting in the baseline. Hence, this is a setting that calls for network-based targeting.

The next three rows of table 4 focus on presenting seed-level explanatory variables of my analysis. I calculate these variables given the seeds chosen by BBMM. In particular, the values represent an average for two seeds, whenever the information on both seed households is available (for 138 villages). Otherwise, it represents the only seed for which the data is available (for 53 villages). To calculate the Eigenvector Centrality of Seeds, I use the eigenvector centrality values that are pre-calculated and available in the BBMM replication dataset.[29] By the design of the experiment, complex seeds have the highest average centrality. BBMM argues that it is due to the optimality of seeding only the most central households when the underlying diffusion process is of complex contagion. Similarly, they expect the simple seeds to have relatively less average centrality than complex seeds as it is optimal to seed one central and one peripheral household when the underlying diffusion process is of simple contagion. BBMM also argue that geo seeds should be less central as they have less than average land by design (which is a measure of less than average wealth), and hence are less likely to be well connected. These patterns are indeed what I observe in the baseline. In terms of the average eigenvector centrality of the seeds, the simple seeds are not statistically different than the benchmark seeds. However, both complex and geo seeds are statically different than the benchmark (at 1% and 5% level of significance, respectively).

Table 4: Baseline Village-level Sample Characteristics

	Treatment Status				
Variable	Benchmark	Complex	Simple	Geo	Overall
Adoption Rate (PP)	0.018	0.030	0.029	0.029	0.026
	(0.035)	(0.063)	(0.060)	(0.077)	(0.060)
Any Non-Seed Adopters (PP)	0.300	0.340	0.320	0.420	0.345
	(0.463)	(0.479)	(0.471)	(0.499)	(0.477)
Eigenvector Centrality of Seeds†	0.178	0.235	0.187	0.129	0.182
	(0.090)	(0.077)	(0.096)	(0.090)	(0.096)
Predicted Adoption Index of Seeds‡	0.110	0.114	0.101	0.082	0.101
	(0.034)	(0.036)	(0.041)	(0.025)	(0.036)
Predicted Usage Index of Seeds‡	0.184	0.186	0.172	0.158	0.175
	(0.031)	(0.032)	(0.042)	(0.024)	(0.035)
CV of Predicted Adoption Index	0.389	0.378	0.379	0.366	0.378
	(0.069)	(0.077)	(0.075)	(0.062)	(0.071)
CV of Predicted Usage Index	0.193	0.188	0.185	0.180	0.187
	(0.039)	(0.040)	(0.037)	(0.033)	(0.038)
Observations	50	50	50	50	200

Notes: † Contains 44 observations for the benchmark treatment group, 49 observations for the other treatment groups. ‡ Contains 48 observations for the complex treatment group. Seed level measures are calculated using the average of two seeds, whenever the information on both seeds are available. Otherwise they reflect the information for one seed. Coefficient of Variations (CV) are calculated at the village level for the whole village. Adoption Rate and Any Non-Seed Adopters are calculated excluding seed or shadow farmers in a village.

I use predicted adoption and usage indices as proxies for the adoption probability. Depending on the proxy, the adoption probabilities differ, but the ranking over different treatment groups remains the same. Here, complex and benchmark seeds have the

[29]Formal definition of eigenvector centrality can be found in Appendix A.

highest adoption probabilities, followed by simple seeds. The geo seeds have the lowest baseline probability of adoption. There are no statistically significant differences between benchmark, complex, and simple seeds. However, geo seeds are statistically different than their benchmark counterparts (at a 1% level of significance).

The final two rows of table 4 present the village-level heterogeneity in adoption probabilities. I measure these using the coefficient of variation (CV) of adoption probability proxies at the village level. In terms of these measures, all other treatment villages have lower heterogeneity in adoption probability than the benchmark villages. However, the geo treatment group is the only one having significantly less heterogeneity than the benchmark group (at the 10% level). The differences are statistically insignificant for complex and simple treatment villages.

Before proceeding to my main empirical results in the following sub-section, let me focus on Figure 8. This figure presents the outcome variables over varying degrees of village-level heterogeneity, where the village-level heterogeneity is proxied by the CV of the predicted adoption index. Here, I categorized the seeding strategy based on the seeds' average centrality and adoption probability. For this figure, I define centrality-based seeds as the seed household(s) that have higher than the median level of average eigenvector centrality at baseline. Similarly, probability-based seed household(s) are defined to have higher than the median level of average adoption probability at baseline. Thus, following this classification, seed household(s) selected in the BBMM experiment can fall under four categories: both centrality-based and probability-based, only centrality-based, only probability-based, and none. Based on my simulations, I expect the effectiveness of centrality-based seeds to decrease as village heterogeneity increases. Similarly, I anticipate the performance of probability-based seeds to improve as village heterogeneity increases. However, I expect these patterns to emerge only in the years 2 and 3 after the interventions. In the first year, after the seeds received training, there was not enough time for diffusion for similar patterns to be evident.[30]

This pattern is what I observe. In years 2 and 3, as village-level heterogeneity increases, the performance of centrality-based seeds decreases compared to their probability-based counterpart. The opposite is true for probability-based seeds compared to the centrality-based seeds. On the contrary, I notice the opposite pattern in year 1 for the adoption rate. However, for the percentage of villages with non-seed adopters, I observe that in the first year the gap between centrality-based and probability-based seeds is closing with an increase in village-level heterogeneity. But, the centrality-based seeds remain the most successful one, for all levels of village heterogeneity.

[30]Training for the seed households took place just a few months before the household survey in year 1. Thus, similar to BBMM, my regression results focus on the effect on the outcome variables from years 2 and 3.

Although informative, the descriptive figures do not account for village-level heterogeneity in terms of other variables. In defining the centrality-based and probability-based seeds as dummy variables, Figure 8 also fail to capture the entire village-level variations of these seeds in terms of their centrality and probability measures. In the following sub-section, I present the reduced form results of my analysis that test my hypotheses more formally.

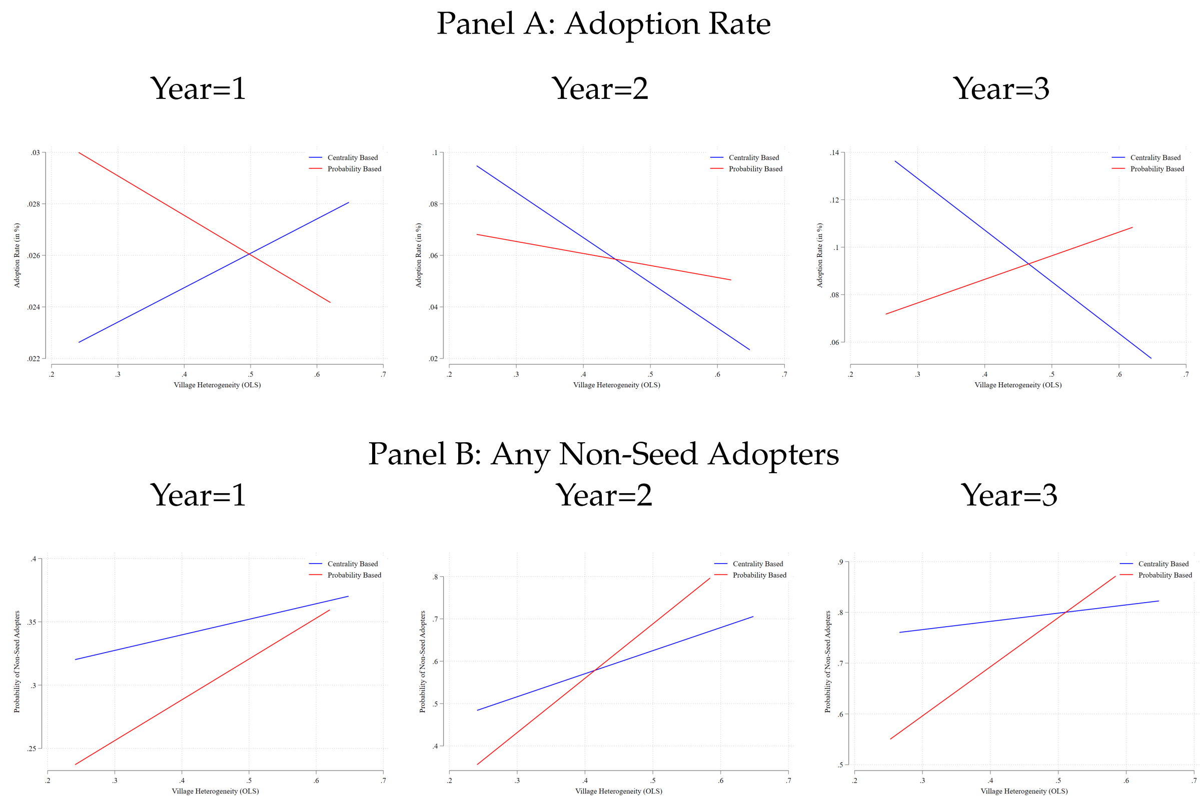


Figure 8: Outcomes for Different Seeding Strategies with respect to Village Heterogeneity

5.3 Reduced Form Results

Table 5 focuses on exploring village-level non-experimental variations. Subsequently, Table 6 presents results exploring the experimental variations. For both these tables, I proxy for the probability of adoption using the predicted adoption index. The results are similar using the predicted usage index as the proxy. Thus, I do not present them here to avoid repetitions. These results are in Appendix D.

For Table 5, the main coefficients of interest are those corresponding to the interactions of $Heterogeneity_v$, with $Centrality_v$ and $Probability_v$. As argued in the last section, I expect the coefficient of $Centrality_v \times Heterogeneity_v$ to be negative and the coefficient of $Probability_v \times Heterogeneity_v$ to be positive. Columns (1) and (2) present the results for the adoption rate, with and without the village level controls. Here, both coefficients of interest are of the desired sign and highly significant. The results show that for a completely homogeneous village, a one standard deviation increase in eigenvector centrality leads to a 9.17%-11.73% increase in the adoption rate. This is a 352.69%-451.15% increase compared to the baseline mean adoption rate of 2.6%. However, for villages with heterogeneity at the level of baseline mean, the effect drops to an increase of only 1.11%-1.81%. Which is a 42.69%-69.62% increase compared to the baseline mean adoption rate. Similarly, one standard deviation increase in predicted adoption decreases the adoption rate by 8.56%-11.89% for a homogeneous village. This is a decrease of 329.23%-457.31%, compared to the mean adoption rate at the baseline. However, for villages with heterogeneity at the level of baseline mean, the effect drops to a decrease of 1.36%-1.74% only. This is a much smaller decrease of 52.31%-66.92% compared to the baseline mean of adoption rate.

Table 5: Village level Regression 1 of Adoption Outcomes (Pit Planting)

Variables	Adoption Rate (1)	(2)	Any Non-Seed Adopters (3)	(4)
Eigenvector Centrality of Seeds	1.173**	0.917*	1.181	1.235
(=$Centrality_v$)	(0.581)	(0.467)	(1.439)	(1.332)
Predicted Adoption Index of Seeds	-2.973**	-2.140	-8.019**	-3.344
(=$Probability_v$)	(1.467)	(1.318)	(3.257)	(3.233)
CV of Predicted Adoption Index	-0.296	-0.157	-0.928	0.506
(=$Heterogeneity_v$)	(0.208)	(0.214)	(1.079)	(1.053)
$Centrality_v \times Heterogeneity_v$	-2.625**	-2.131**	-2.851	-3.299
	(1.324)	(1.066)	(3.777)	(3.562)
$Probability_v \times Heterogeneity_v$	6.715**	4.762*	18.480***	7.562
	(3.131)	(2.796)	(6.997)	(7.073)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.080	0.180	0.049	0.169

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses. All regressions include a constant term and year fixed effects. Village-level controls include percentage of village using pit planting at baseline, percentage of village using compost at baseline, percentage of village using fertilizer at baseline, village size, the square of village size, and district fixed effects.

The results for Any Non-Seed Adopters are in columns (3) and (4), with and without the village level controls. Although the coefficients of interest are of the desired sign, they are mostly insignificant. The results show that for completely homogeneous villages, one standard deviation increase in eigenvector centrality leads to an 11.81%-12.35% increase in the probability of having at least one non-seed adopter. Compared to the baseline mean of 34.5% for the variable Any Non-Seed Adopters, this is an increase of 34.23%-35.80%. However, for villages with heterogeneity at the level of baseline mean, the effect drops to being between a 0.12% decrease and a 1.03% increase. This is between a drop of 0.35% and an increase of 2.99%, compared to the baseline mean of the dependent variable. On the other hand, one standard deviation increase in predicted adoption decreases the probability of having at least one non-seed adopter by 13.38%-32.08% for a homogeneous village. Compared to the baseline mean of the variable Any Non-Seed Adopters, this is a decrease of 38.78%-92.99%. For villages with heterogeneity at the level of baseline mean, however, the effect drops to a probability decrease of 1.94%-4.13% only. Which is a drop of only 5.62%-11.97% compared to the baseline mean of the dependent variable.

Table 6 focuses on exploring between treatment group variations. Here, I am interested in the coefficients of $Centrality_v \times Heterogeneity_v$ and $Probability_v \times Heterogeneity_v$, across different treatment groups. Note that the sign of $Centrality_v \times Heterogeneity_v$ is negative, and the sign of $Probability_v \times Heterogeneity_v$ is positive within different treatment groups, in line with the results of table 5. Table 6 notes the differences in the coefficients of $Centrality_v \times Heterogeneity_v$ and $Probability_v \times Heterogeneity_v$, across different treatment groups. Some of these differences are statistically significant, while others are not. However, the signs are all consistent with my discussion in the last section.

Columns (5) and (6) present the results for Adoption Rate, with and without the village level controls. The results show that for a completely homogeneous village, one standard deviation increase in the centrality of seed households leads to a 6.33%-7.75% improvement in the adoption rate. This is a 243.46%-298.08% increase compared to the baseline mean adoption rate of 2.6%. However, for benchmark villages having heterogeneity at the level of baseline benchmark mean, the effect drops to a decrease of 1.68%-2.37%. Which is a 64.62%-91.15% decrease compared to the baseline mean adoption rate. The negative effect of heterogeneity on the relationship between seeds' centrality and the adoption rate is lower for the other treatment groups compared to the benchmark. However, the difference is statistically significant only for the complex and geo treatment groups. Similarly, one standard deviation increase in the adoption probability of seed households decreases the adoption rate by 6.32%-9.45% for a homogeneous village, which is a decrease of 243.08%-363.46% compared to the mean adoption rate at the baseline. However, for benchmark villages having heterogeneity at the level of baseline benchmark mean, the effect drops between a decrease of 0.30% and an increase of 0.07%. This is a much smaller effect of between 11.54% decrease and 2.69% increase, compared to the baseline mean of adoption rate. The positive impact of heterogeneity on the relationship between seeds' adoption probability and the adoption rate is lower for complex and simple treatment groups. However, these differences are statistically insignificant. Only for the geo treatment group, the impact is significantly lower compared to the benchmark.

The results for Any Non-Seed Adopters are in columns (7) and (8), with and without the village level controls. For this outcome variable, the effects are also in the same direction for all the treatment groups.

Table 6: Village level Regression 2 of Adoption Outcomes (Pit Planting)

Variables	Adoption Rate (5)	(6)	Any Non-Seed Adopters (7)	(8)
Eigenvector Centrality of Seeds	0.775*	0.633*	1.703	1.638
(= $Centrality_v$)	(0.423)	(0.378)	(1.660)	(1.468)
Predicted Adoption Index of Seeds	-2.362**	-1.578	-10.42***	-5.947*
(= $Probability_v$)	(1.091)	(1.024)	(3.679)	(3.566)
CV of Predicted Adoption Index	-0.321	-0.150	-0.923	0.417
(= $Heterogeneity_v$)	(0.206)	(0.200)	(1.105)	(1.073)
$Centrality_v \times Heterogeneity_v$	-2.423**	-2.237**	-6.692	-6.574
	(1.093)	(0.996)	(4.503)	(4.119)
$Centrality_v \times Heterogeneity_v \times Complex$	0.657**	0.664**	4.328**	3.756**
	(0.306)	(0.282)	(1.775)	(1.664)
$Centrality_v \times Heterogeneity_v \times Simple$	0.416	0.428	1.078	0.431
	(0.337)	(0.320)	(2.060)	(1.947)
$Centrality_v \times Heterogeneity_v \times Geo$	2.026**	1.942**	0.103	-0.0702
	(0.940)	(0.839)	(2.235)	(2.098)
$Probability_v \times Heterogeneity_v$	5.881**	4.104*	22.97***	12.35
	(2.437)	(2.286)	(7.720)	(7.626)
$Probability_v \times Heterogeneity_v \times Complex$	-0.155	-0.232	-1.275	-0.679
	(0.520)	(0.497)	(2.765)	(2.654)
$Probability_v \times Heterogeneity_v \times Simple$	-0.121	-0.110	1.941	3.511
	(0.642)	(0.571)	(3.572)	(3.333)
$Probability_v \times Heterogeneity_v \times Geo$	-2.588**	-2.562**	-0.391	0.538
	(1.131)	(1.039)	(4.028)	(3.618)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.133	0.224	0.113	0.222

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses. All regressions include a constant term and year fixed effects. Village-level controls include percentage of village using pit planting at baseline, percentage of village using compost at baseline, percentage of village using fertilizer at baseline, village size, the square of village size, and district fixed effects.

The results show that for homogeneous villages, one standard deviation increase in eigenvector centrality leads to a 16.38%-17.04% improvement in the probability of having at least one non-seed adopter. Compared to the baseline mean of 34.5% for the variable Any Non-Seed Adopters, this is an increase of 47.48%-49.39%. But, for benchmark villages having heterogeneity at the level of baseline benchmark mean, the effect drops to around a 9% decrease of probability, which is a 26.09% drop compared to the baseline mean of the dependent variable. The negative impact of heterogeneity on the relationship between seeds' centrality and the probability of having non-seed adopters is lower for the simple treatment group. It is different for the geo treatment group as well. However, only for the complex treatment group, the effect is significantly lower compared to the benchmark. On the other hand, one standard deviation increase in the predicted adoption index decreases the probability of having at least one non-seed adopter by 23.79%-41.68% for a homogeneous village. Compared to the baseline mean of the variable Any Non-Seed Adopters, this is a decrease of 68.96%-120.81%. However, for benchmark villages having heterogeneity at the level of baseline benchmark mean, the effect drops to a decrease of 4.57%-5.94%. This is a drop of only 13.25%-17.22% compared to the baseline mean of the dependent variable. The positive impact of heterogeneity on the relationship between seeds' adoption probability and the probability of having non-seed adopters is lower for the complex treatment group, higher for the simple treatment group, and different for the geo treatment group. However, none of these differences are statistically significant.

These results show that for homogeneous villages seeding central households leads to improvements in adoption. Existing literature recognizes the role played by central agents in improving diffusion, and subsequent adoption of a product. BBMM uses the same data to show that more central seeds cause higher adoption. Seeds' centrality is one of the main reasons for improved adoption of a microfinance product in India by Banerjee et al. (2013), and improved take-up of an insurance product in China by Cai et al. (2015). I add to this literature by providing evidence that the positive effect of seeds' centrality decreases as the target population becomes more heterogeneous. In addition, I show evidence in favor of an alternative probability-based seeding strategy to work better in such a scenario.

6 Summary and Concluding Remarks

I focus on network-based targeting strategies for improving technology adoption when a new technology has more benefits to some agents than others. This heterogeneity in benefits can be due to the agents differing in terms of their education, skills, and ability, which affect how much they can learn about the new technology and use it in practice. We can also attribute the heterogeneity to the agents' input choices and their access to other technologies. In particular, I assume that this heterogeneity in benefits directly impacts the diffusion of information regarding the benefits of the new technology. This assumption deviates from the existing literature that considers information diffusion to depend on existing social ties only. I present a model that helps formalize such a scenario, adding to the theoretical literature that considers households homogeneous in what they need to learn about new technologies. I use simulations, building on the structure of my theoretical model, to generate testable hypotheses on the performance of different network-based targeting strategies. I hypothesize that the relative performance of different targeting strategies depends on the population heterogeneity in terms of the expected benefits of adopting a technology. In particular, I expect centrality-based targeting to perform worse as the heterogeneity increase but targeting based on the adoption probability to perform better if the network is highly assortative in terms of characteristics determining the benefits. To test these hypotheses, I use the replication data of BBMM collected from Malawi. To generate variation in the BBMM sample in the benefits of a new technology, I use the AESTAS dataset also collected from Malawi. Reduced form results lend support in favor of my hypotheses. Exploring non-experimental village-level variations, I show that the positive effect of the seed households' centrality on the adoption of pit planting decrease with an increase in village-level heterogeneity in terms of adoption probability. Simultaneously, the negative effect of the seed households' adoption probability decreases with an increase in village-level heterogeneity. Although weaker, I find similar results when I shift my focus to exploring the experimental variations of BBMM.

The reduced form analysis uses a series of assumptions that I discuss in detail. The main challenge in targeting based on the adoption probability is that the adoption probabilities depend on the benefits realized only after the adoption. I attempt to solve this issue by using additional data to predict adoption probability conditional on observable demographics. A better approach would be to collect more information on the same households making the adoption decisions. For that purpose, along with better identification, a randomized controlled trial that mimics the ideal experiment discussed in this paper would be more suitable. A randomized controlled trial of such a nature could also help me disentangle the effects of centrality and probability of seeds. These, along with a more structural approach can help separately identify the performances of targeting strategies discussed here. These are exciting avenues for future research.

For policy, my results suggest that network-based targeting may require more than identifying central households within a social network. More specifically, I argue for the need to understand possible population heterogeneity in benefits. This recommendation adds to the existing literature that highlights the importance of central agents for targeting policies (Beaman et al. (2021a)) and focuses on cost-effectively identifying these agents (Banerjee et al. (2019)). This recommendation is applicable only if a new technology is such that there can be sufficient population heterogeneity in terms of its benefits. In practice, this demands more information than the requirement for just identifying central households, increasing the cost of network-based targeting. This increase in the cost of network-based targeting may make it more attractive to randomly select more seed households following the approach proposed by Akbarpour et al.

(2021). We need a proper cost-benefit analysis for this purpose, which is beyond the scope of this paper.

References

ACEMOGLU, D., A. NEDIC, AND A. OZDAGLAR (2008): "Convergence of rule-of-thumb learning rules in social networks," in *2008 47th IEEE Conference on Decision and Control*, 1714–1720.

ACEMOGLU, D., A. OZDAGLAR, AND E. YILDIZ (2011): "Diffusion of innovations in social networks," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, 2329–2334.

AKBARPOUR, M., S. MALLADI, AND A. SABERI (2021): "Just a Few Seeds More: Value of Network Information for Diffusion," *Unpublished Manuscript*.

BALA, V. AND S. GOYAL (1998): "Learning from Neighbours," *Review of Economic Studies*, 65, 595–621.

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

BANERJEE, A., E. BREZA, A. G. CHANDRASEKHAR, AND B. GOLUB (2018): "When Less is More: Experimental Evidence on Information Delivery During India's Demonetization," Working Paper 24679, National Bureau of Economic Research.

BANERJEE, A., E. BREZA, A. G. CHANDRASEKHAR, AND M. MOBIUS (2021): "Naïve Learning with Uninformed Agents," *American Economic Review*, 111, 3540–74.

BANERJEE, A., A. G. CHANDRASEKHAR, E. DUFLO, AND M. O. JACKSON (2013): "The Diffusion of Microfinance," *Science*, 341, 1236498.

——— (2019): "Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials," *The Review of Economic Studies*, 86, 2453–2490.

BARNETT-HOWELL, Z. AND A. M. MOBARAK (2021): "Chapter 84 - Social networks analysis in agricultural economies," in *Handbook of Agricultural Economics*, ed. by C. B. Barrett and D. R. Just, Elsevier, vol. 5 of *Handbook of Agricultural Economics*, 4613–4652.

BEAMAN, L., A. BENYISHAY, J. MAGRUDER, AND A. M. MOBARAK (2021a): "Can Network Theory-Based Targeting Increase Technology Adoption?" *American Economic Review*, 111, 1918–43.

——— (2021b): "Data and Code for: Can Network Theory-based Targeting Increase Technology Adoption?" *Nashville, TN: American Economic Association*.

BENYISHAY, A. AND A. M. MOBARAK (2018): "Social Learning and Incentives for Experimentation and Communication," *The Review of Economic Studies*, 86, 976–1009.

BESLEY, T. AND A. CASE (1993): "Modeling Technology Adoption in Developing Countries," *The American Economic Review*, 83, 396–402.

BLOCH, F., M. O. JACKSON, AND P. TEBALDI (2021): "Centrality measures in networks," *Available at SSRN 2749124*.

BOLD, T., K. C. KAIZZI, J. SVENSSON, AND D. YANAGIZAWA-DROTT (2017): "Lemon Technologies and Adoption: Measurement, Theory and Evidence from Agricultural Markets in Uganda," *The Quarterly Journal of Economics*, 132, 1055–1100.

BREZA, E., A. CHANDRASEKHAR, B. GOLUB, AND A. PARVATHANENI (2019): "Networks in economic development," *Oxford Review of Economic Policy*, 35, 678–721.

BUSTOS, P., B. CAPRETTINI, AND J. PONTICELLI (2016): "Agricultural Productivity and Structural Transformation: Evidence from Brazil," *American Economic Review*, 106, 1320–1365.

CAI, J., A. DE JANVRY, AND E. SADOULET (2015): "Social Networks and the Decision to Insure," *American Economic Journal: Applied Economics*, 7, 81–108.

CATALINI, C. AND C. TUCKER (2017): "When early adopters don't adopt," *Science*, 357, 135–136.

CHANDRASEKHAR, A. G., B. GOLUB, AND H. YANG (2018): "Signaling, Shame, and Silence in Social Learning," Working Paper 25169, National Bureau of Economic Research.

CHANDRASEKHAR, A. G., H. LARREGUY, AND J. P. XANDRI (2020): "Testing models of social learning on networks: Evidence from two experiments," *Econometrica*, 88, 1–32.

CHENG, H. W. J. (2021): "Factors Affecting Technological Diffusion Through Social Networks: A Review of the Empirical Evidence," *The World Bank Research Observer*, lkab010.

CONLEY, T. G. AND C. R. UDRY (2010): "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, 100, 35–69.

DEGROOT, M. H. (1974): "Reaching a Consensus," *Journal of the American Statistical Association*, 69, 118–121.

DEMARZO, P. M., D. VAYANOS, AND J. ZWIEBEL (2003): "Persuasion Bias, Social Influence, and Unidimensional Opinions*," *The Quarterly Journal of Economics*, 118, 909–968.

FEDER, G. AND S. SAVASTANO (2006): "The role of opinion leaders in the diffusion of new knowledge: The case of integrated pest management," *World Development*, 34, 1287–1300.

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

GOLUB, B. AND M. O. JACKSON (2010): "Naïve Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, 2, 112–49.

IFPRI (2021a): "Agricultural Extension Services and Technology Adoption Survey, 2016," *Harvard Dataverse*.

——— (2021b): "Agricultural Extension Services and Technology Adoption Survey, 2018," *Harvard Dataverse*.

JACKSON, M. O. (2010): *Social and Economic Networks*, Princeton University Press.

KHAILA, S., F. TCHUWA, S. FRANZEL, AND B. SIMPSON (2015): "The Farmer-to-Farmer Extension Approach in Malawi: A Survey of Lead Farmers," ICRAF Working Paper 189, World Agroforestry Centre.

KRISHNAN, P. AND M. PATNAM (2013): "Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption?" *American Journal of Agricultural Economics*, 96, 308–327.

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

MAGRUDER, J. R. (2018): "An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries," *Annual Review of Resource Economics*, 10, 299–316.

MILLER, G. AND A. M. MOBARAK (2015): "Learning About New Technologies Through Social Networks: Experimental Evidence on Nontraditional Stoves in Bangladesh," *Marketing Science*, 34, 480–499.

MUNSHI, K. (2004): "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution," *Journal of Development Economics*, 73, 185–213.

——— (2007): "Chapter 48 Information Networks in Dynamic Agrarian Economies," in *Handbook of Development Economics*, ed. by T. P. Schultz and J. A. Strauss, Elsevier, vol. 4, 3085–3113.

NEWMAN, M. (2010): *Networks: An Introduction*, Oxford University Press.

NIU, C. AND C. RAGASA (2018): "Selective attention and information loss in the lab-to-farm knowledge chain: The case of Malawian agricultural extension programs," *Agricultural Systems*, 165, 147–163.

RAGASA, C. (2020): "Effectiveness of the lead farmer approach in agricultural extension service provision: Nationally representative panel data analysis in Malawi," *Land Use Policy*, 99, 104966.

RAGASA, C., D. MZUNGU, K. KALAGHO, AND C. KAZEMBE (2021): "Impact of interactive radio programming on agricultural technology adoption and crop diversification in Malawi," *Journal of Development Effectiveness*, 13, 204–223.

RAGASA, C. AND C. NIU (2017): "The state of agricultural extension and advisory services provision in Malawi: Insights from household and community surveys," Tech. rep., International Food Policy Research Institute (IFPRI).

SURI, T. (2011): "Selection and Comparative Advantage in Technology Adoption," *Econometrica*, 79, 159–209.

TJERNSTRÖM, E. (2017): "Learning from others in heterogeneous environments," *Unpublished Manuscript*.

Appendices

A Mathematical Definitions

The objective of this section is to formally define the network centrality measures used in different parts of this paper. This section heavily draws from Chapter 2 of Jackson (2010) and Chapter 7 of Newman (2010). More detailed descriptions along with some applications can be found in these sources.

Let $N = \{1, 2, ..., n\}$ be a set of agents (called nodes) involved in a network. The tuple (N, g) defines a *graph* (or, *network*), where g is a real-valued $n \times n$ matrix (called adjacency matrix) with g_{ij} representing the (possibly) weighted and/or directed relation between i and j. [31] An *edge* (i, j) is defined as a link from i to j.[32] Edge (i, j) exists if and only if $g_{ij} \neq 0$. A sequence of edges $(i_1, i_2), (i_2, i_3), ..., (i_{k-1}, i_k)$ is called a *walk*. A *path* between i and j is defined as a walk such that $i_1 = i$ and $i_k = j$, with each node being distinct in the walk. A *geodesic path* between two nodes i and j is defined as a path with no more edges than any other paths between these nodes. In other words, geodesic path(s) between i and j represent(s) the shortest distance from i to j.[33]

Degree Centrality: For an undirected and unweighted network (N, g), degree centrality of a node k is given by:

$$\mathscr{D}_k(N, g) = \sum_{i=1}^{n} g_{ki},$$

which measures the number of nodes connected with node k. For a directed and unweighted network (N, g), nodes have both in-degree and out-degree. Out-degree of node k measures the number of nodes the node k connects to:

$$\mathscr{D}_k^{out}(N, g) = \sum_{i=1}^{n} g_{ki}.$$

[31]Networks can be either weighted or unweighted. For a unweighted network, g_{ij} is either 0 or 1 representing whether i is connected to j or not. For weighted network, g_{ij} can take other non-negative values. The weights represent the intensity of relationships. Networks can also be either directed or undirected. For a directed network, I define g_{ij} to be representing a link from i to j, and g_{ji} to be representing a link from j to i. In an undirected network, $g_{ij} = g_{ji}\ \forall i, j \in N$. Alternatively, in a directed network, $\exists i, j \in N,\ s.t.\ g_{ij} \neq g_{ji}$. For my theoretical model, I consider networks to be unweighted and undirected. In the BBMM experiment, the networks were considered to be weighted and undirected.

[32]Which is the same as the edge (j, i) in an undirected network. Same may not be true for a directed network.

[33]The calculation uses weights associated with the edges in the path(s).

Similarly, in-degree of node k measures the number of nodes connected to node k:

$$\mathscr{D}_k^{in}(N,g) = \sum_{i=1}^{n} g_{ik}.$$

For weighted networks, the same measure is termed as the *strength* of node k.

Betweenness Centrality: Let P_{ij}^k denote the number of geodesic paths from i to j that passes through k, with P_{ij} being the total number of geodesic paths from i to j. Then the betweenness centrality of node k in network (N,g) is defined to be:

$$\mathscr{B}_k(N,g) = \sum_{\forall i,j \ s.t. i \neq j \ and \ k \notin \{i,j\}} \left(\frac{P_{ij}^k}{P_{ij}}\right),$$

with $\frac{P_{ij}^k}{P_{ij}} = 0$ if $P_{ij} = 0$.

Closeness Centrality: Let L_{ki} denote the number of edges in the shortest path between k and i. Then the closeness centrality of node k in network (N,g) is defined as:

$$\mathscr{C}_k(N,g) = \frac{(n-1)}{\sum_{i \neq k} L_{ki}}.$$

For an undirected graph, we consider distances between k and every other node. Alternatively, for a directed graph, the distances from every other node to k is considered.

Eigenvector Centrality: For an undirected network (N,g), the eigenvector centrality $\mathscr{E}_k(N,g)$ of node k is defined as:

$$\lambda \mathscr{E}_k(N,g) = \sum_{\forall i} g_{ki} \mathscr{E}_i(N,g),$$

where $\mathscr{E}(N,g) = \{\mathscr{E}_1(N,g), \mathscr{E}_2(N,g), ..., \mathscr{E}_N(N,g)\}$ is an eigenvector of g with λ being the corresponding eigenvalue. It is conventional to use the eigenvector associated with the largest eigenvalue.

For a directed network (N,g), the adjacency matrix g is asymmetric. So, there are two sets of eigenvectors: left eigenvectors (uses the connection of each nodes to other nodes) and right eigenvectors (use the connection of other nodes to each nodes). Conventionally, the right eigenvector is considered to be more important, which is a measure of how many other nodes are pointing towards a node. For a directed network (N,g), the right-eigenvector centrality $\mathscr{E}_k^R(N,g)$ of node k can be defined as:

$$\lambda_R \mathscr{E}_k^R(N,g) = \sum_{\forall i} g_{ik} \mathscr{E}_i^R(N,g),$$

where $\mathscr{E}^R(N,g) = \{\mathscr{E}_1^R(N,g), \mathscr{E}_2^R(N,g), ..., \mathscr{E}_N^R(N,g)\}$ is a right-eigenvector of g with

λ^R being the corresponding eigenvalue. Again, it is conventional to use the eigenvector associated with the largest eigenvalue.

It is important to note that a node having only outgoing edges will have a right eigenvector centrality of zero in a directed network. The same is true for any node that has incoming edges only from nodes that have only outgoing edges. In general, any node whose all incoming connections can be traced back to node(s) with only outgoing edges will have a right eigenvector centrality of zero in a directed network. This is a problematic property for eigenvector centrality in a directed network. Since I consider only undirected networks, I do not need to worry about this.

B Details on the Simulation Method

The following 6 steps were taken for the simulations used in this paper:

Step 1: Generating true probabilities of success

The first step is to generate p^*_{iH}s for a network. For that purpose, I first randomly set p^0_{iH}s to be either 0 or 1. Then use the following process to generate p^*_{iH}s:

$$p^*_{iH} = \sum_{j \in \mathcal{I}} M^T_{ij} p^0_{jH},$$

where T is the number of iterations and M is an $n \times n$ weighted and non-negative matrix such that $\sum_{j \in \mathcal{I}} M_{ij} = 1$.

If I set T to be large enough, it would lead all p^*_{iH}s to converge to an average. On the other hand, lower values of T keep p^*_{iH}s heterogeneous. So, I can vary T to control the degree of heterogeneity in terms of p^*_{iH}s. The elements of M is drawn from a normal distribution $N(-4, 4)$, which was first truncated so that all negative values get replaced by 0 and then normalized such that each row sums up to 1. The normalization may not work if for row k of the truncated matrix, the elements sum up to 0. In such cases, all elements of the row k were first replaced with a 0. Then, M_{kk} was set to be 1. Once the matrix M is prepared, p^*_{iH} are constructed for different values of T. This procedure is repeated independently 200 times for each T, to generate 200 villages with differing levels of heterogeneity in terms of p^*_{iH}s, independent from each other.

Step 2: Generating networks of households

Once the p^*_{iH}s are generated, the next task is to generate networks assorted in terms of these p^*_{iH}s. For that purpose, I generate adjacency matrix g such that $g_{ij} = 1$ if $|p^*_{iH} - p^*_{jH}| < 0.1$, 0 otherwise.[34] I then generate the influence matrix G by normalizing each row of g. Remember that $G_{ij} \geq 0$ represents the weight i places on j's opinion (with $\sum_{j \in \mathcal{I}} G_{ij} = 1$ and $G_{ii} \neq 0$). This procedure is repeated to generate 200 village

[34] Note that this ensures $g_{ii} = 1$.

networks for each value of T.

Step 3: Selecting seeding strategy

Once I generate 200 villages with corresponding G and p^*_{iH}s, the next step is to study the effectiveness of different seeding strategies. For a given network, I consider the initial beliefs to be equal to 0 for all households: $\hat{p}^H_{i0} = 0$. The seeded households then get informed. Consider node k to be one of the seeds, then I exogenously set $\hat{p}^H_{k0} = p^*_{kH}$. I choose two seed households per village, in line with the experimental framework of BBMM. The policy question is: which two households should be chose in a given village? I consider two different targeting strategies:

- **Centrality-based:** Choose two households that have maximum average centrality in a network.
- **Probability-based:** Choose two households that have maximum average p^*_{iH}s in a network.

Step 4: Diffusion

Given the seeding strategy in a network, I let the diffusion take place for 10 periods. In each of these periods, each uninformed node (the nodes that do not know their p^*_{iH}s) makes a decision of whether or not to get informed based on their $\hat{p}^H_{it}$. For that, each period t, they compare $\hat{p}^H_{it}$ with a threshold $\bar{p}^H_i := \bar{p}^*_{iH} + \bar{\eta}_i$. I set the threshold $\bar{p}^H_i = \bar{p}^H = 0.5$, for all households in different networks. If for any period t, $\hat{p}^H_{it} > \bar{p}^H$, the household is considered informed next period onward ($\hat{p}^H_{is} = p^*_{iH} \forall s > t$).

Step 5: Evaluation

Given the value of parameter T, in a set of 200 networks, I evaluate the targeting efficiency on average. Targeting efficiency of strategy κ is measured by the following equation in each network:

$$Efficiency_\kappa = \frac{\Psi_\kappa}{\Psi}.$$

Here Ψ denotes the number of informed households under full efficiency. On the other hand, Ψ_κ captures the number of households that end up getting informed within 10 periods of implementing the targeting strategy κ, among those that should get informed.

Step 6: Comparison

The evaluation is done for both centrality-based and probability-based seeding strategies, for varying degree of heterogeneity in p^*_{iH}s. As discussed above, this heterogeneity is controlled by the parameter T. The results are then compared, across varying degree of heterogeneity in p^*_{iH}s, between centrality-based and probability-based targeting.

C Construction of Adoption and Usage Indices

To calculate the adoption index in the AESTAS data, I use the self-reported adoption for a list of pre-determined technologies and practices. This includes the following 13 agricultural practices:

1. Soil cover
2. Zero or minimum tillage
3. Crop rotation
4. Intercropping
5. Crop residue incorporation
6. Composting pits or piles
7. Composting toilets
8. Agroforestry
9. Bunds or ridges
10. Pit planting
11. Planting vetivar grass
12. Water harvesting in pits or swales or dug outs
13. Manure or fertilizer making

As well as the following 5 food processing practices:

1. Including multiple food groups (dietary diversity) in each meal
2. Consuming iron-rich foods
3. Using iodized salt in food preparation
4. Washing hands before preparing and consuming food
5. Food, health and nutrition

The adoption variables are available in the data as a set of dummy variables (1 implies adoption, 0 implies no adoption). I take the average of these set of 18 dummy variables to calculate the adoption index.

To calculate the usage index, I use the self-reported plot-level usage for the following list of 19 agricultural technologies:

1. Contour bunds
2. Box ridges
3. Field leveling
4. Soil cover
5. Mulching
6. Zero or minimum tillage
7. Plowing with power tiller or animal tractor
8. Herbicide before planting
9. Herbicide after planting
10. Transplanting the seedlings
11. Rain water harvesting, water retention or water management practice
12. Proper plant spacing
13. Pesticide
14. Putting crop residue on top of the soil (without soil disturbance)
15. Crop residue incorporation (with soil disturbance)
16. Getting soil sample to have it tested by soil experts
17. Asking advice from plant clinic or plant doctors
18. Pit planting
19. Row planting

The usage variables are available in the data, for both dry and rainy seasons, as a set of dummy variables (1 implies usage, 0 implies no usage). First, I take the max of these dummy variable per technology, for each year. Then I take the average of a set of 19 dummy variables to calculate the usage index.

D Robustness Checks

D.1 Simulations

Table D.7: Simulation Robustness (w.r.t different centrality measure)

Strategy	Statistic	Level of Correlation Low	Medium	High
Betweenness Centrality-Based	Mean	0.38	0.35	0.42
	Variance	0.10	0.12	0.16
Probability-Based	Mean	0.46	0.38	0.33
	Variance	0.03	0.07	0.11
	Observations	200	200	178

Notes: Simulations on varying levels of correlation are all done for 200 networks, each containing 20 households. However, upon generation of the true probabilities, some networks are dropped as they contained 0% of informed households under full efficiency.

D.2 Empirical Results

Table D.8: OLS Results for Adoption and Usage (Pooled vs. Individual Years)

	Adoption Index			Usage Index		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Adults	0.008***	0.009***	0.006**	0.011***	0.014***	0.008***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Children	0.003***	0.004**	0.003**	0.003***	0.001	0.005***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Housing	0.009***	0.013***	0.005*	0.003	0.003	0.003
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Livestock	0.010***	0.014***	0.007*	0.014***	0.020***	0.007**
	(0.003)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
Assets	0.024***	0.014***	0.034***	0.020***	0.011***	0.029***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Year	Pooled	2016	2018	Pooled	2016	2018
Observations	5610	2803	2805	5610	2803	2805
R-squared	0.096	0.082	0.125	0.085	0.088	0.103

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors clustered at the section level are in parentheses. All regressions use sample weights and include a constant term. Household controls are not included. The variables *Adults* and *Children* represent number of adults and children in a household, respectively. The variables *Housing*, *Livestock*, and *Assets* were standardized first principal components.

Table D.9: Non-Linear Regression Results for Adoption and Usage

	Tobit				Negative Binomial			
	Adoption Index		Usage Index		Adoption Sum		Usage Sum	
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adults	0.016***	0.012***	0.012***	0.009***	0.103***	0.079***	0.061***	0.045***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.026)	(0.026)	(0.010)	(0.010)
Children	0.006***	0.004**	0.004***	0.003***	0.042***	0.026**	0.021***	0.014***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.012)	(0.011)	(0.005)	(0.005)
Housing	0.016***	0.015***	0.003	0.002	0.101***	0.101***	0.020	0.017
	(0.004)	(0.004)	(0.002)	(0.002)	(0.024)	(0.024)	(0.012)	(0.012)
Livestock	0.016***	0.007**	0.015***	0.010***	0.104***	0.049**	0.067***	0.044***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.024)	(0.023)	(0.011)	(0.010)
Assets	0.048***	0.035***	0.022***	0.016***	0.294***	0.216***	0.127***	0.089***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.024)	(0.025)	(0.012)	(0.013)
Baseline Mean	0.084	0.084	0.138	0.138	1.510	1.510	2.615	2.615
(Standard Deviation)	(0.123)	(0.123)	(0.115)	(0.115)	(2.215)	(2.215)	(2.192)	(2.192)
Household Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	5610	5606	5610	5606	5610	5606	5610	5606
pseudo R-squared	0.246	0.357	-0.107	-0.155	0.027	0.039	0.020	0.029

Notes: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Robust standard errors clustered at the section level are in parentheses. All regressions use sample weights and include a constant term. The variables *Adults* and *Children* represent number of adults and children in a household, respectively. The variables *Housing*, *Livestock*, and *Assets* were standardized first principal components. Household Controls include: gender and age of household head, activity of household head (0- Non-Farmer, 1- Farmer), whether the household applied for a loan in the past, the households' time and risk preferences, and whether a household member is a lead farmer (LF).

Table D.10: Robustness of Village level Regression 1 with respect to different set of controls

	Adoption Rate				Any Non-Seed Adopters			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eigenvector Centrality of Seeds	1.173**	1.250*	0.917*	0.981*	1.181	1.150	1.235	1.210
(= $Centrality_v$)	(0.581)	(0.635)	(0.467)	(0.517)	(1.439)	(1.477)	(1.332)	(1.396)
Predicted Adoption Index of Seeds	-2.973**	-2.880**	-2.140	-2.087*	-8.019**	-8.645**	-3.344	-3.832
(= $Probability_v$)	(1.467)	(1.331)	(1.318)	(1.226)	(3.257)	(3.379)	(3.233)	(3.337)
CV of Predicted Adoption Index	-0.296	-0.223	-0.157	-0.093	-0.928	-0.806	0.506	0.669
(= $Heterogeneity_v$)	(0.208)	(0.184)	(0.214)	(0.194)	(1.079)	(1.108)	(1.053)	(1.096)
$Centrality_v \times Heterogeneity_v$	-2.625**	-2.857**	-2.131**	-2.365**	-2.851	-3.636	-3.299	-4.218
	(1.324)	(1.407)	(1.066)	(1.158)	(3.777)	(3.835)	(3.562)	(3.714)
$Probability_v \times Heterogeneity_v$	6.715**	6.628**	4.762*	4.779*	18.48***	19.67***	7.562	8.921
	(3.131)	(2.912)	(2.796)	(2.644)	(6.997)	(7.126)	(7.073)	(7.197)
Village-level Controls	No	No	Yes	Yes	No	No	Yes	Yes
Treatment Dummies	No	Yes	No	Yes	No	Yes	No	Yes
Observations	324	324	324	324	324	324	324	324
R-squared	0.080	0.092	0.180	0.190	0.049	0.094	0.169	0.210

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses. All regressions include a constant term and year fixed effects. Village-level controls include percentage of village using pit planting at baseline, percentage of village using compost at baseline, percentage of village using fertilizer at baseline, village size, the square of village size, and district fixed effects.

Table D.11: Robustness of Village level Regression 2 with respect to different set of controls

	Adoption Rate				Any Non-Seed Adopters			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eigenvector Centrality of Seeds	0.775*	0.950*	0.633*	0.755	1.703	3.655**	1.638	3.101**
(= $Centrality_v$)	(0.423)	(0.518)	(0.378)	(0.469)	(1.660)	(1.795)	(1.468)	(1.561)
Predicted Adoption Index of Seeds	-2.362**	-1.786	-1.578	-1.114	-10.420***	-6.356	-5.947*	-2.864
(= $Probability_v$)	(1.091)	(1.139)	(1.024)	(1.076)	(3.679)	(3.990)	(3.566)	(3.828)
CV of Predicted Adoption Index	-0.321	-0.102	-0.150	0.016	-0.923	1.070	0.417	1.955*
(= $Heterogeneity_v$)	(0.206)	(0.219)	(0.200)	(0.206)	(1.105)	(1.229)	(1.073)	(1.161)
$Centrality_v \times Heterogeneity_v$	-2.423**	-2.784**	-2.237**	-2.488**	-6.692	-10.790**	-6.574	-9.434**
	(1.093)	(1.365)	(0.996)	(1.249)	(4.503)	(5.001)	(4.119)	(4.464)
$Centrality_v \times Heterogeneity_v \times Complex$	0.657**	0.821**	0.664**	0.816**	4.328**	4.361*	3.756**	3.329
	(0.306)	(0.362)	(0.282)	(0.356)	(1.775)	(2.369)	(1.664)	(2.192)
$Centrality_v \times Heterogeneity_v \times Simple$	0.416	0.369	0.428	0.390	1.078	0.775	0.431	0.122
	(0.337)	(0.389)	(0.320)	(0.362)	(2.060)	(2.285)	(1.947)	(2.138)
$Centrality_v \times Heterogeneity_v \times Geo$	2.026**	1.685*	1.942**	1.668**	0.103	-3.716	-0.070	-3.651
	(0.940)	(0.903)	(0.839)	(0.794)	(2.235)	(2.297)	(2.098)	(2.280)
$Probability_v \times Heterogeneity_v$	5.881**	4.908**	4.104*	3.359	22.970***	16.590**	12.350	8.010
	(2.437)	(2.488)	(2.286)	(2.345)	(7.720)	(8.021)	(7.626)	(7.785)
$Probability_v \times Heterogeneity_v \times Complex$	-0.155	-0.176	-0.232	-0.274	-1.275	-2.498	-0.679	-2.400
	(0.520)	(0.594)	(0.497)	(0.567)	(2.765)	(2.649)	(2.654)	(2.612)
$Probability_v \times Heterogeneity_v \times Simple$	-0.121	-0.623	-0.110	-0.624	1.941 -0.097	3.511	1.274	
	(0.642)	(0.779)	(0.571)	(0.761)	(3.572)	(4.231)	(3.333)	(4.091)
$Probability_v \times Heterogeneity_v \times Geo$	-2.588**	-4.335**	-2.562**	-3.952**	-0.391	-18.540***	0.538	-14.550***
	(1.131)	(1.719)	(1.039)	(1.669)	(4.028)	(5.667)	(3.618)	(5.240)
Village-level Controls	No	No	Yes	Yes	No	No	Yes	Yes
Treatment Dummies	No	Yes	No	Yes	No	Yes	No	Yes
Observations	324	324	324	324	324	324	324	324
R-squared	0.133	0.141	0.224	0.229	0.113	0.154	0.222	0.249

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses. All regressions include a constant term and year fixed effects. Village-level controls include percentage of village using pit planting at baseline, percentage of village using compost at baseline, percentage of village using fertilizer at baseline, village size, the square of village size, and district fixed effects.

Table D.12: Village level Regression 1 with Different Measure of Probability

Variables	Adoption Rate (1)	(2)	Any Non-Seed Adopters (3)	(4)
Eigenvector Centrality of Seeds	0.999*	0.817*	0.984	1.067
(=$Centrality_v$)	(0.565)	(0.480)	(1.302)	(1.191)
Predicted Usage Index of Seeds	-2.174	-1.511	-4.599	-0.084
(=$Probability_v$)	(1.410)	(1.279)	(3.317)	(3.053)
CV of Predicted Usage Index	-1.091	-0.631	-2.549	2.142
(=$Heterogeneity_v$)	(0.805)	(0.779)	(2.905)	(2.823)
$Centrality_v \times Heterogeneity_v$	-4.481*	-3.936*	-4.874	-5.907
	(2.623)	(2.281)	(6.889)	(6.438)
$Probability_v \times Heterogeneity_v$	10.330*	7.276	23.130	0.889
	(6.160)	(5.623)	(14.190)	(13.400)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.063	0.174	0.037	0.164

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses. All regressions include a constant term and year fixed effects. Village-level controls include percentage of village using pit planting at baseline, percentage of village using compost at baseline, percentage of village using fertilizer at baseline, village size, the square of village size, and district fixed effects.

Table D.13: Village level Regression 2 with Different Measure of Probability

Variables	Adoption Rate (5)	(6)	Any Non-Seed Adopters (7)	(8)
Eigenvector Centrality of Seeds	0.730	0.644	1.525	1.482
(=*Centrality*$_v$)	(0.471)	(0.446)	(1.528)	(1.337)
Predicted Usage Index of Seeds	-1.975	-1.400	-7.027*	-2.854
(=*Probability*$_v$)	(1.200)	(1.148)	(3.982)	(3.619)
CV of Predicted Usage Index	-1.203	-0.727	-3.640	0.546
(=*Heterogeneity*$_v$)	(0.755)	(0.731)	(3.203)	(3.116)
$Centrality_v \times Heterogeneity_v$	-4.619*	-4.617*	-12.420	-12.190
	(2.549)	(2.473)	(8.555)	(7.660)
$Centrality_v \times Heterogeneity_v \times Complex$	1.432*	1.595**	9.431**	8.099**
	(0.749)	(0.720)	(4.323)	(3.996)
$Centrality_v \times Heterogeneity_v \times Simple$	0.492	0.576	3.308	1.958
	(0.860)	(0.831)	(4.665)	(4.340)
$Centrality_v \times Heterogeneity_v \times Geo$	3.957*	3.711**	-1.692	-2.661
	(2.057)	(1.785)	(4.676)	(4.495)
$Probability_v \times Heterogeneity_v$	10.260*	7.702	33.700*	13.410
	(5.561)	(5.378)	(17.390)	(16.260)
$Probability_v \times Heterogeneity_v \times Complex$	-0.316	-0.589	-2.606	-1.839
	(0.762)	(0.778)	(4.577)	(4.315)
$Probability_v \times Heterogeneity_v \times Simple$	0.428	0.416	1.355	3.119
	(0.984)	(0.866)	(5.269)	(4.868)
$Probability_v \times Heterogeneity_v \times Geo$	-2.468*	-2.409**	2.565	3.786
	(1.377)	(1.217)	(4.925)	(4.505)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.114	0.212	0.100	0.215

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses. All regressions include a constant term and year fixed effects. Village-level controls include percentage of village using pit planting at baseline, percentage of village using compost at baseline, percentage of village using fertilizer at baseline, village size, the square of village size, and district fixed effects.

Table D.14: Village level Regression 1 with Different Measure of Centrality

Variables	Adoption Rate (1)	(2)	Any Non-Seed Adopters (3)	(4)
Closeness Centrality of Seeds	0.609**	0.454*	0.571	0.617
(=*$Centrality_v$*)	(0.306)	(0.234)	(0.709)	(0.659)
Predicted Adoption Index of Seeds	-2.438**	-1.709	-7.555**	-2.904
(=*$Probability_v$*)	(1.230)	(1.134)	(3.201)	(3.152)
CV of Predicted Adoption Index	-0.077	-0.007	-0.677	0.887
(=*$Heterogeneity_v$*)	(0.214)	(0.202)	(1.196)	(1.158)
$Centrality_v \times Heterogeneity_v$	-1.325*	-1.020*	-1.552	-1.997
	(0.716)	(0.558)	(1.896)	(1.823)
$Probability_v \times Heterogeneity_v$	5.610**	3.814	17.550**	6.849
	(2.660)	(2.439)	(6.873)	(6.940)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.087	0.179	0.048	0.170

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses. All regressions include a constant term and year fixed effects. Village-level controls include percentage of village using pit planting at baseline, percentage of village using compost at baseline, percentage of village using fertilizer at baseline, village size, the square of village size, and district fixed effects.

Table D.15: Village level Regression 2 with Different Measure of Centrality

Variables	Adoption Rate (5)	(6)	Any Non-Seed Adopters (7)	(8)
Closeness Centrality of Seeds	0.497**	0.336*	0.603	0.727
(=$Centrality_v$)	(0.242)	(0.183)	(0.713)	(0.707)
Predicted Adoption Index of Seeds	-1.734	-1.077	-9.416**	-5.382
(=$Probability_v$)	(1.056)	(0.986)	(3.663)	(3.520)
CV of Predicted Adoption Index	0.001	0.059	-0.627	0.912
(=$Heterogeneity_v$)	(0.216)	(0.213)	(1.228)	(1.205)
$Centrality_v \times Heterogeneity_v$	-1.457**	-1.181**	-2.508	-3.114
	(0.591)	(0.478)	(1.935)	(1.939)
$Centrality_v \times Heterogeneity_v \times Complex$	0.307**	0.304**	1.446*	1.355*
	(0.137)	(0.140)	(0.838)	(0.810)
$Centrality_v \times Heterogeneity_v \times Simple$	0.364**	0.395***	-0.401	-0.498
	(0.157)	(0.152)	(0.934)	(0.917)
$Centrality_v \times Heterogeneity_v \times Geo$	0.679**	0.667**	0.517	0.140
	(0.267)	(0.262)	(0.988)	(0.914)
$Probability_v \times Heterogeneity_v$	4.791**	3.306	19.310***	9.942
	(2.281)	(2.166)	(7.105)	(6.963)
$Probability_v \times Heterogeneity_v \times Complex$	-0.351	-0.419	0.056	0.189
	(0.632)	(0.637)	(3.155)	(3.031)
$Probability_v \times Heterogeneity_v \times Simple$	-1.125*	-1.235*	4.299	5.406
	(0.664)	(0.629)	(3.876)	(3.727)
$Probability_v \times Heterogeneity_v \times Geo$	-2.855**	-2.864**	-2.748	0.060
	(1.200)	(1.187)	(4.867)	(4.398)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.121	0.209	0.109	0.223

Notes: * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses. All regressions include a constant term and year fixed effects. Village-level controls include percentage of village using pit planting at baseline, percentage of village using compost at baseline, percentage of village using fertilizer at baseline, village size, the square of village size, and district fixed effects.