# Network-Based Targeting with Heterogeneous Agents for Improving Technology Adoption

This version: March 14, 2024. Click here for the latest version.

Aranya Chakraborty*

**Abstract**

Can we use social ties to improve technology adoption? I examine this question when the benefits from a new technology vary in the population, with such heterogeneity affecting the diffusion process. I develop a theoretical framework of information diffusion in a network that exhibits the possibility of low information equilibria where agents sub-optimally decide not to adopt new technology, highlighting the need for interventions for information diffusion. My simulations suggest that the optimal network-based intervention in such a scenario relies on the underlying heterogeneity in the population. More importantly, network-centrality-based interventions recommended in the literature fail to be effective, and an alternative adoption probability-based intervention may work better under some conditions. I test these predictions using data from Malawi and provide evidence supporting the theoretical model. My results suggest that population heterogeneity in benefits from a technology affects the success of alternative network-based interventions that promote that technology.

**JEL Codes:** D83, O13, O33, Q16.

**Keywords:** Targeting, Social Network, Technology Adoption, Agriculture.

---

*Amrut Mody School of Management, Ahmedabad University. M-315, School of Arts and Sciences Building, Ahmedabad University Gate-10, Commerce Six Road, Navrangpura, Ahmedabad 380009, Gujarat, India. E-mail: aranya.chakraborty@ahduni.edu.in.
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, Hassan Benchekroun, Munir Squires, Nicolas Gendron-Carrier, Fernando Saltiel, Mallesh Pai, Juni Singh, and Adam Aberra for their helpful comments. I would also like to extend my gratitude to the participants at the CITD-JNU International Conference, the 18th Annual Conference on Economic Growth and Development at ISI-Delhi, the North East Universities Development Consortium (NEUDC) Conference 2023, the McGill-CIREQ Lunch Seminars, McGill CPD Social Statistics and Population Dynamics Seminar, McGill Economics Applied Micro Breakfasts, the 56th Annual Conference of the Canadian Economics Association (CEA), and the 17th CIREQ Ph.D. Students' Conference for their generous feedback. I am also thankful to Hassan Benchekroun for giving me the opportunity to present and get feedback from the students of the Ph.D. (Economics) Research Seminar course at McGill University. 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. Declarations of interest: none.

# 1 Introduction

Technology adoption in agriculture drives 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 a phenomenon (Magruder, 2018). How can 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. Given a threshold-based characterization of the information diffusion process, 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). However, the recommendation assumes 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 via experimentation about a new technology. Since experimentation 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 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 the 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. The benefits are only realized after adoption, so factoring them into the targeting strategies is difficult. I attempt to solve this issue using additional data to estimate adoption conditional on observable demographics. This way, I can categorize the population's 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[3] to improve the adoption of a technology suitable for my analysis, thus including all other information that I need. I exploit the village-level and experimental variations in the BBMM data to test my hypotheses.

[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. 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.

[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 adoption patterns 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 the network literature, information entry points are called *seeds*.

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 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 the benefits. The intuition behind such a result lies in the characteristics of the central seeds in a network. 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 regarding 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 specific sub-section of the population, targeting based on centrality is more likely to fail in reaching the population of interest. Targeting the population of interest works better in such a scenario.

My empirical results show evidence in favor of my hypothesis. I show that the positive effect of seeds' centrality on the adoption of pit planting decreases with an increase in village-level heterogeneity in the probability of adoption. Simultaneously, the negative impact of seeds' probability of adoption drops with increased village-level heterogeneity.

The study makes three contributions to the existing literature. First, I provide theoretical and empirical evidence 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, several studies have focused 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

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

targeting policies (e.g., Beaman et al., 2021a; Akbarpour et al., 2021). However, these studies implicitly assume existing network ties are the only factor characterizing diffusion. 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 directly affecting 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.[6]

Second, my theoretical framework helps formalize the scenario where agents learn from their network about a technology more beneficial to some of them than others. Existing theories on the diffusion of information regarding a technology in a network consider technologies equally helpful to everyone. The adoption may still differ due to heterogeneity in costs. However, these heterogeneous costs are mostly assumed to be known by the agents and thus do not require learning.[7] Thus, 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

[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 about using fertilizers for pineapple production in Ghana from their veteran neighbors.

[6] Although de Janvry et al. (2022) document the consequences of village-level heterogeneity on farmer-level adoption decisions, they do not focus on the consequences of their findings on the design of network-based targeting strategies.

[7] For heterogeneous costs unknown to the agents, there is usually learning via experimentation but no social learning.

(Suri, 2011). For example, the performance of some agricultural practices may depend on land quality.[8] Thus, the benefits of some technologies may vary depending on the agent-specific characteristics (Crane-Droesch, 2017). The consequences of this heterogeneity on the diffusion of knowledge have not previously been modeled 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 the network is highly assortative in terms of the heterogeneity.[9] Meanwhile, 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) shows that communicators who share a group identity with the farmers or face comparable agricultural conditions better convince farmers to adopt a new technology. Feder and Savastano (2006) takes a middle ground in arguing that the most influential 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 offers the simulations that help me form the hypotheses for this study. Section 4 discusses the hypotheses, my empirical strategy for testing them, and the data I use. Section 5 presents and discusses my empirical results. Finally, in Section 6, I summarize my findings and make concluding remarks.

[8]In Munshi (2004), adopting 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.

[9]I define early adopters as households 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).

# 2 Theoretical Framework

I consider a choice problem that requires learning in a social network. The problem involves the adoption of a technology when agents vary in terms of the benefits they get from the 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 make the adoption decision only after learning about the benefits via experimentation. As experimentation is costly, agents rely on their social ties to determine whether or not to experiment.

Similar to Golub and Jackson (2010), I consider agents to have an initial opinion and be involved in DeGroot learning (developed in DeGroot (1974) and DeMarzo et al. (2003)).[10] 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.[11] In addition, I consider the possibility that agents are heterogeneous regarding their distribution of payoffs associated with the new technology.

## 2.1 Perfect Information Benchmark

Consider a traditional technology that has a sure payoff of $\pi^T$ and new technology that provides a payoff of $\pi^N(\omega_{it})$, where $\omega_{it} \in \Omega$ is the state of the world parameter drawn independently at each period $t$ according to the distribution $p_i^*(\omega_{it})$ for household $i$.[12] To simplify, further 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 $\omega_H$), and the other where the new technology has a lower payoff than the traditional one (denoted

[10]DeGroot learning is considered as it is used in all the canonical models of information aggregation in the development literature. There is empirical evidence in favor of it (see Corazzini et al., 2012 and Chandrasekhar et al., 2020).

[11]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 have an initial opinion (this includes partially informed agents) and informed agents are fully informed.

[12]The assumption of draws not being correlated over time within a household helps me abstract away from the problem where households observe the draws over time and update their beliefs accordingly. The assumption of draws not being correlated over time between households constrains how the households can learn from each other.

$\omega_L$). Let $p^*_{iH} := p^*_i(\omega_H)$ denote the probability that for household $i$ the new technology has a higher payoff than the traditional one.

Assuming households to be risk-neutral and myopic, the households' *technology-adoption decision* is then to compare the net benefit of adopting the new technology with that of the traditional technology.[13] The benefit of adopting the new technology for household $i$ is:

$$\sum_{s\in\{H,L\}} p^*_{is}\pi^N(\omega_s) = p^*_{iH}\pi^N(\omega_H) + p^*_{iL}\pi^N(\omega_L), \text{ where } p^*_{iL} := p^*_i(\omega_L) = 1 - p^*_{iH} \tag{1}$$

Let $c_i$ denote the cost of adopting the new technology for household $i$,[14] then the household should adopt the new technology if and only if:

$$\sum_{s\in\{H,L\}} p^*_{is}\pi^N(\omega_s) - c_i \geq \pi^T. \tag{2}$$

As I have now defined the necessary notations, let me first formally state the assumption made above regarding the states of the world:

**Assumption 1:** *$\forall it$, $\exists \omega_{it}, \omega'_{it} \in \Omega$ 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.*

For the simplified case of $\Omega = \{\omega_H, \omega_L\}$, this gives us the condition: $\pi^N(\omega_H) \geq \pi^T \geq \pi^N(\omega_L)$, which is how I defined $\omega_H$ and $\omega_L$ above. This is a necessary assumption to ensure

[13]The assumption of risk neutrality is for simplification purposes only, as it allows us to focus solely on the expected values without considering the variation around them. As the new technology is assumed to be riskier than the traditional technology here, risk-averse 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. One can easily accommodate this 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 myopia assumption is also necessary for simplification purposes. It helps me focus on a static model instead of a dynamic one. Moreover, in a social learning model, non-myopic households may strategically wait for their peers to learn before deciding to learn. This will lead to sequential social learning as defined in Golub and Sadler (2016), where agents will wait for their peers to be informed first, which is beyond the scope of this paper.

[14]I assume the adoption costs to be heterogeneous and privately known to the agents and focus on the uncertainty surrounding the unknown expected benefits in the following subsection. This is a simplified scenario, which can easily extend to cases where costs are uncertain, too. Similarly, one can extend the social learning described in the following subsection to model learning about partially correlated unknown heterogeneous costs.

that one technology doesn't dominate the other in terms of benefits in all states of the world. I further assume:

<u>**Assumption 2:**</u> *$\exists i, j \in \mathcal{I}$ such that $\sum_{s\in\{H,L\}} p^*_{is}\pi^N(\omega_s) - c_i > \pi^T$ and $\sum_{s\in\{H,L\}} p^*_{js}\pi^N(\omega_s) - c_j < \pi^T$, where $\mathcal{I}$ denote the set of all households.*

The assumption implies that there is enough heterogeneity in the population so that some households get positive net benefits from adopting the new technology instead of the traditional one, while others do not. This assumption ensures that the new technology is *better* for only a fraction of households in the population, making social learning *noisy* in this context.

## 2.2 Imperfect Information and Learning

So far, I have assumed households to have perfect information regarding $p^*_{iH}$. Next, I relax that assumption and allow for the possibility that households can be uninformed about $p^*_{iH}$. Let $p^H_{it}$ be household $i$'s belief of $p^*_{iH}$ at period $t$.[15] The beliefs can be informed or uninformed. For informed households, $p^H_{it} = p^*_{iH}$. On the contrary, the uninformed households need to put effort into experimentation $e_{it} \in \{0, 1\}$ to learn $p^*_{iH}$.[16] If $e_{it} = 1$, the household $i$ at period $t$ learns about $p^*_{iH}$ at cost $\eta_i$.[17] I additionally assume that households incur the cost of experimentation only once - the first time they get informed via experimentation, and once they are informed, they remain informed forever.[18]

Before learning about $p^*_{iH}$ via costly experimentation, the households can use DeGroot averaging to approximate $p^*_{iH}$ costlessly with information from their social ties. Let $G$ denote the $n \times n$ weighted and non-negative influence matrix ($n = |\mathcal{I}|$), where $G_{ij} \geq 0$

[15]Similarly, $p^L_{it} = 1 - p^H_{it}$ is household $i$'s belief of $p^*_{iL}$ at period $t$.

[16]I make no assumptions on the initial number of informed households. 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.

[17]I simplify the experimentation stage to be a one-shot process where agents invest in a costly effort and learn about $p^*_{iH}$ with certainty. This is to abstract away from the *learning from experimentation* and focus on the *learning from doing*. One can extend this model to the situation where investment in experimentation leads to learning about $p^*_{iH}$ with some uncertainty, which is beyond the scope of my analysis. Also, note that the heterogeneity in $p^*_{iH}$ and the myopic nature of the households stop them from free-riding on each others' experimentation in this scenario.

[18]In other words, if we think of $e_{it} = 1$ as an indicator of the household $i$ being informed at period $t$, then if $e_{i\tau} = 1$, then $e_{it} = 1$ $\forall t \geq \tau$.

represents the weight $i$ places on $j$'s opinion (with $\sum_{j \in \mathcal{I}} G_{ij} = 1$ and $G_{ii} \neq 0$).[19] Then $\hat{p}^H_{it} = \sum_{j \in \mathcal{I}} G_{ij} p^H_{jt-1}$ denotes household $i$'s approximation based on their aggregation of opinions following the DeGroot averaging. This brings me to the next assumption:

**Assumption 3:** *The networks are assortative in $p^*_{iH}$s, i.e., $G_{ij} \neq 0$ if $|p^*_{iH} - p^*_{jH}| < \delta$, where $\delta$ is a small number.*[20]

The rationale behind such an assumption is twofold. First, it is well-recognized in the technology adoption literature that connected agents share similar characteristics that help them benefit similarly from a technology (Munshi, 2007). For example, focusing only on the geographic connections, it is easy to argue that neighboring farmers share soil quality. As a result, they benefit similarly from an agricultural technology whose outcomes depend on soil quality. For network connections, the similarity extends beyond geographic characteristics. I expect households to sort according to their socio-economic attributes. As the benefits from any technology depend on these socio-economic attributes, connected households sharing socio-economic attributes should benefit similarly from the technology.

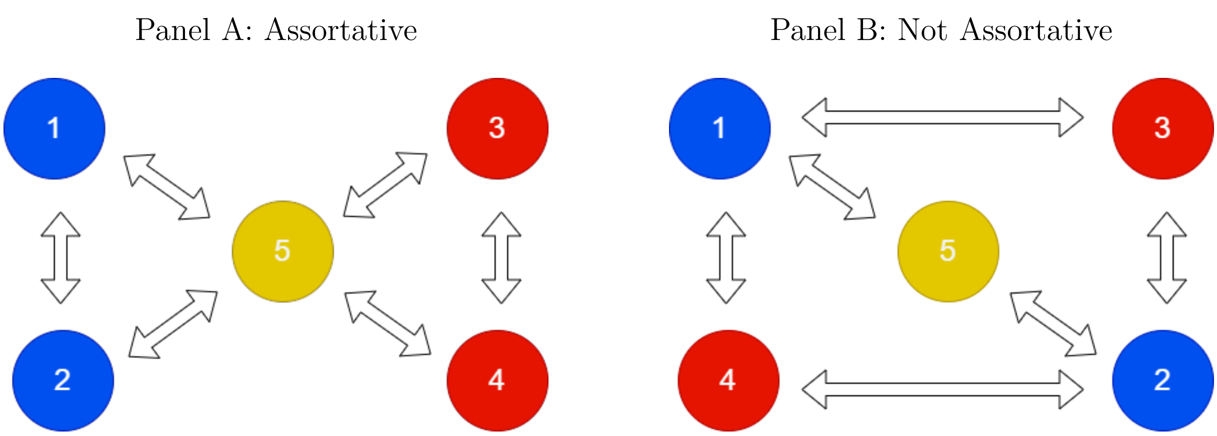


Figure 1: Networks with Heterogeneous Benefits

In addition, the *assortative* property is necessary for the social ties to be informative with varying $p^*_{iH}$s. To demonstrate this, consider Figure 1. The panels of Figure 1 present

[19] I make no further assumptions regarding the $G_{ij}$s for the theoretical model. However, in the simulations presented in the next section, they are assumed to be equal whenever they are non-zero.

[20] The assortativity property is similar to the *homophily* defined in Golub and Jackson (2012a,b,c). In their model, homophily is the '*tendency of agents to associate disproportionately with those having similar traits*.' I consider assortativity to be a more specific version of homophily, where the agents are linked only with those that share similar traits and *not* with others. This leads to the scenario where agents place positive weight only on those with similar characteristics. One can extend the model to the more general scenario of homophily, where the agents disproportionately weigh their connections based on their similarity in characteristics with those connections.

heterogeneous networks, with the colors representing benefits from some technology. In both figures, Agents numbered 1 and 2 should not adopt the technology as their benefits are low (represented by the color blue). Similarly, Agents 3 and 4 should adopt the technology as their benefits are high (represented by red). Finally, Agent number 5 would benefit moderately from the technology (represented by yellow). The figures differ, however, in the network ties (given by the arrows).

Panel A presents an assortative network. Here, uninformed agents can use their social ties to form a belief close to their true types. To see this, consider Agent 4, who should adopt the technology if informed. If uninformed, she would seek information from Agents 3 and 5. Agents 3 and 5 observe weakly higher than average benefits from the technology if they are informed. Thus, informed agents 3 and 5 can influence agent 4 in making the right choice regarding whether or not to put effort into experimentation to learn about the new technology. Contrast this with the network in Panel B, which is a non-assortative network. Here, if uninformed, Agent 4 would seek information from Agents 1 and 2. Agents 1 and 2 observe lower-than-average benefits from the technology if informed. Thus, if Agent 4 is uninformed and seeks information from her informed network ties, she will be influenced to make the wrong choice regarding whether or not to put effort into experimentation to learn about the new technology. Thus, social learning will not help agents make the right choice for the network in Panel B.[21]

I assume a two-stage decision-making process of technology adoption that builds on Chandrasekhar et al. (2018), Banerjee et al. (2021), and BBMM. The timeline of decision-making in the model is as follows:

1. Every period, uninformed household $i$ decides whether or not to put effort into experimentation to learn about $p^*_{iH}$.[22]

2. To decide, they collect information on beliefs $p^H_{jt-1}$ from their peers $j \in \mathcal{I}$, formed in the last period (some informed, some uninformed) and use DeGroot averaging to calculate $\hat{p}^H_{it} = \sum_{j \in \mathcal{I}} G_{ij} p^H_{jt-1}$.

[21] It is worth noting that if we assume $p^*_{iH} = p^*_H \; \forall i \in \mathcal{I}$ (i.e., assume homogeneity in success probabilities) the network ties become automatically helpful in making the right choice.

[22] Informed households already know their $p^*_{iH}$, so they do not need to make this decision.

3. Based on $\hat{p}_{it}^H$, then they decide whether or not to put effort into experimentation to get informed. The following rule represents this choice:

$$e_{it} = \begin{cases} 1 \; if \; \sum_{s\in\{H,L\}} \hat{p}_{it}^s \pi^N(\omega_s) - c_i - \pi^T \geq \eta_i \\ 0 \text{ otherwise.} \end{cases} \tag{3}$$

4. If they do not get informed ($e_{it} = 0$), their new belief is formed to be equal to the DeGroot average ($p_{it}^H = \hat{p}_{it}^H$), and next period they repeat from the step 1 above. On the other hand, if they get informed ($e_{it} = 1$), they now know $p_{iH}^*$ and make adoption decisions based on that, with $p_{is}^H = p_{iH}^* \; \forall s \geq t$. This implies that in any period $t$, we can represent household $i$'s belief as:

$$p_{it}^H = e_{it} p_{iH}^* + (1 - e_{it}) \hat{p}_{it}^H. \tag{4}$$

Informed households make the adoption decision following the rule:

$$\text{Adopt}_{it} = \begin{cases} 1 \; if \; \sum_{s\in\{H,L\}} p_{is}^* \pi^N(\omega_s) - c_i \geq \pi^T \\ 0 \text{ otherwise.} \end{cases} \tag{5}$$

## 2.3 Implications

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:

$$\begin{aligned} \sum_{s\in\{H,L\}} p_{is}^* \pi^N(\omega_s) - c_i \geq \pi^T &\Rightarrow p_{iH}^* \pi^N(\omega_H) + (1 - p_{iH}^*) \pi^N(\omega_L) - c_i \geq \pi^T \\ &\Rightarrow p_{iH}^* \geq \frac{c_i + (\pi^T - \pi^N(\omega_L))}{(\pi^N(\omega_H) - \pi^N(\omega_L))} =: \bar{p}_{iH}^*. \end{aligned} \tag{6}$$

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:

$$
\begin{aligned}
&\sum_{s\in\{H,L\}} \hat{p}^s_{it}\pi^N(\omega_s) - c_i - \pi^T \geq \eta_i \\
&\Rightarrow \hat{p}^H_{it}\pi^N(\omega_H) + (1-\hat{p}^H_{it})\pi^N(\omega_L) - c_i - \pi^T \geq \eta_i \\
&\Rightarrow \hat{p}^H_{it} \geq \frac{c_i + (\pi^T - \pi^N(\omega_L))}{(\pi^N(\omega_H) - \pi^N(\omega_L))} + \frac{\eta_i}{(\pi^N(\omega_H) - \pi^N(\omega_L))} =: \bar{p}^*_{iH} + \bar{\eta}_i. \qquad (7)
\end{aligned}
$$

The condition (7) considers the cost of effort ($\eta_i$). This is because the decision in step 1 is regarding whether or not to put effort into experimentation to be informed. From (6) and (7), 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 (6) is not satisfied, then (7) should not hold if the diffusion of information is efficient. In other words, under fully efficient information diffusion, only those adopting the technology in step 2 would get informed in step 1. Thus, for these households, the following condition must hold:

$$
p^*_{iH} \geq \bar{p}^*_{iH} + \bar{\eta}_i. \qquad (8)
$$

Equation (8) 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 not get 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 experimentation ($\eta_j$). This feature is similar to the models of Chandrasekhar et al. (2018) and Banerjee et al. (2018), where the social stigma of information-seeking stops some 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 setup will help household $i$ make the right decision regarding putting effort into experimentation. The problem, however, arises when most households are uninformed. The situation is particularly interesting when $p_{it}^{H} \approx 0 \ \forall it$. 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 interventions can help in such a scenario. We can exogeneously give information to some households (seeds) to improve adoption. If household $i$ gets exogenously informed about their $p_{iH}^{*}$ at period $t$, household $j$ will update their $\hat{p}_{jt+1}^{H}$ if $j$ puts positive weight on $i$'s opinion. Subsequently, this will lead household $k$ to update their $\hat{p}_{kt+2}^{H}$ if $k$ puts positive weight on $j$'s opinion, and so on. The outcome of this intervention, a few periods down the line regarding the experimentation decision (and eventually the technology adoption decision), will depend on the initial seeding 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 us understand the relative effectiveness of different targeting strategies.

In the next section, I measure the relative performance of two types of such targeting strategies (as compared to a random seeding strategy) using simulations. In doing so, I consider the household networks facing the decision problem described in this section. I focus on the scenarios where initially $p_{it}^{H} \approx 0 \ \forall it$, 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 modeled in the last section with initial $p_{it}^H \approx 0 \; \forall it$. I first demonstrate the potential problem for a centrality-based seeding strategy, recommended in the literature, with the example of a specific network. Then, I simulate multiple networks to analyze whether the problem persists on average under different underlying assumptions and compare the centrality-based seeding strategy with a probability-based seeding strategy (defined below). I show that the relative performance of different targeting strategies depends on the population's heterogeneity level in benefits and whether the networks are assortative in this heterogeneity.

## 3.1 An Illustrative Example

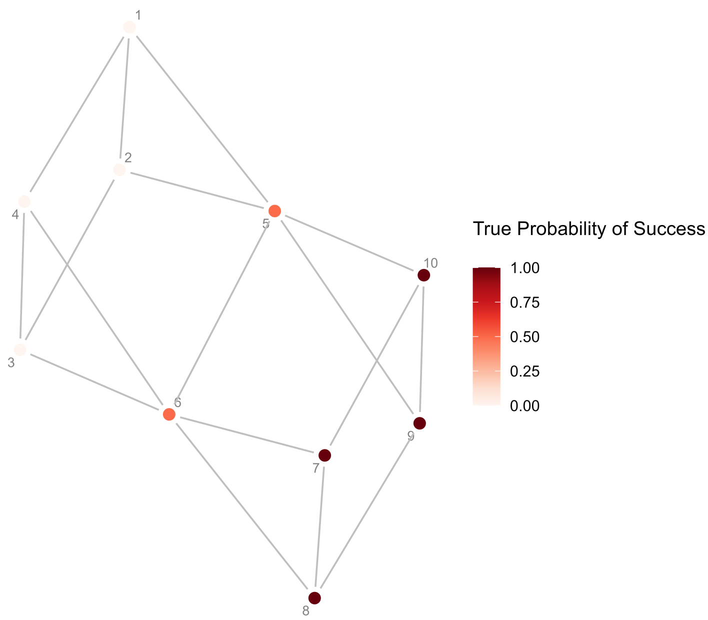


Figure 2: Distribution of True Probability within the network

I start with the example of a specific network of 10 households, depicted in Figure 2. The households are heterogeneous concerning their $p_{iH}^*$s (represented by the nodes' colors), and the network is perfectly assortative in the $p_{iH}^*$s. The network has three types of households: those with high $p_{iH}^*$ (represented by dark red nodes, numbered 7-10), low $p_{iH}^*$ (represented by white nodes, numbered 1-4), and moderate $p_{iH}^*$ (represented by light red nodes, numbered 5 and 6). For this example, consider the threshold probability of learning (i.e. $(\bar{p}^*_{iH} + \bar{\eta}_i)$ in (7)) to be 0.25 for every household. Thus, if $p^*_{iH}$ of a household is more than 25%, the household should get informed if they make their optimal choice. Given the distribution of $p^*_{iH}$s 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 probability of success with the new technology is zero ($p^H_{it} = 0 \; \forall it$). Under such a scenario, even if it is optimal for some households to adopt the technology, they do not. An intervention is then required to improve adoption. The objective behind such an intervention is to ensure that the households that would have adopted the technology under perfect information decide to adopt it. At the same time, for efficiency, we want to ensure that the households that should not adopt the technology under perfect information optimally choose not to put effort into experimentation to get informed about it. Thus, we can measure the efficiency of a seeding strategy $\kappa$ as:

$$\text{Efficiency}_\kappa = \underbrace{\frac{\text{Informed}^T_\kappa}{\text{Informed}^T}}_{A_\kappa} - \underbrace{\frac{\text{Informed}^F_\kappa}{\text{Uninformed}^T}}_{B_\kappa} \tag{9}$$

Here $\textit{Informed}^T$ denotes the number of non-seed households that should have put effort into experimentation to get informed, as they would have adopted the technology under perfect information (i.e., they satisfy equation (8)). Additionally, $\textit{Informed}^T_\kappa$ captures the number of non-seed households that get informed within some periods of implementing the targeting strategy $\kappa$, among those households in $\textit{Informed}^T$. Thus, the term $A_\kappa$ represents the informed non-seed households as a fraction of non-seed households that should have gotten informed under perfect information, given the targeting strategy $\kappa$. Thus, a higher value of this fraction indicates a more successful targeting strategy. The term $B_\kappa$, on the other hand, represents the fraction of non-seed households that are mistargeted by the targeting strategy $\kappa$. $\textit{Uninformed}^T$ denotes the number of non-seed households that should not put effort into experimentation to get informed and $\textit{Informed}^F_\kappa$ captures the number of non-seed households that end up getting informed among those households given targeting strategy $\kappa$. Thus, a higher value of $B_\kappa$ indicates a less successful targeting strategy.

To understand the terms in the efficiency measure more clearly, consider the network in Figure 2. If households numbered 5 and 6 are the seeds for the strategy $\kappa$, we are interested in understanding the efficiency of $\kappa$ for diffusing knowledge among the other households. As four other households (numbered 7-10) should put effort into experimentation to get informed under optimal choices, $\textit{Informed}^T = 4$. Similarly, the other four households (numbered 1-4) should not put effort into experimentation to get informed under optimal choices, implying $\textit{Uninformed}^T = 4$. Now, consider the scenario where the households numbered 1, 8, 9, and 10 decided to put effort into experimentation for getting informed, given the same seeds for strategy $\kappa$. If that is true, then $\textit{Informed}^T_\kappa = 3$ since 8, 9, and 10 are among the households that should get informed. On the other hand, $\textit{Informed}^F_\kappa = 1$ as the household number 1 got mistargeted. In this case, we would have $A_\kappa = 3/4$, $B_\kappa = 1/4$, and $\textit{Efficiency}_\kappa = 3/4 - 1/4 = 1/2$.

My analysis focuses 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 supports in favor (Banerjee et al., 2013), and because BBMM recommends centrality-based targeting for the diffusion process described here.[23] I consider probability-based targeting as an alternative to centrality-based targeting. The probability-based targeting strategy is to seed households with the highest expected benefits with the new technology (i.e., the highest $p^*_{iH}$s in the network). These households are more likely to adopt a technology with a homogeneous cost of experimentation for everyone. Hence, we can think of them as 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.

[23]The diffusion process described in this paper falls under the category of *complex diffusion*. Complex diffusion models assume that information diffuses to an agent 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).

Whereas the centrality-based approach relies on households 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 the average. Second, there is a debate in the existing literature regarding whether opinion leaders should be somewhat *superior* to their followers for the 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. 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 G for detailed results). Eigenvector-based centrality measures represent a household's connectivity to other households, considering the importance of their connections in terms of their respective connections in a recursive manner. A formal definition of different centrality measures can be found in Appendix A.[24] 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 $p^*_{iH}$ 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 exogenously informed about their $p^*_{iH}$ in the first period of policy intervention. 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 links each. Additionally, we can observe that both households have a moderate value for $p^*_{iH}$. This feature is not surprising given that central households are the

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

most connected in the network, and the network is highly assortative in the $p^*_{iH}$s. Thus, the central households represent the average $p^*_{iH}$s in the network, not the ones with a high $p^*_{iH}$. This feature has consequences for the final performance of this targeting strategy.

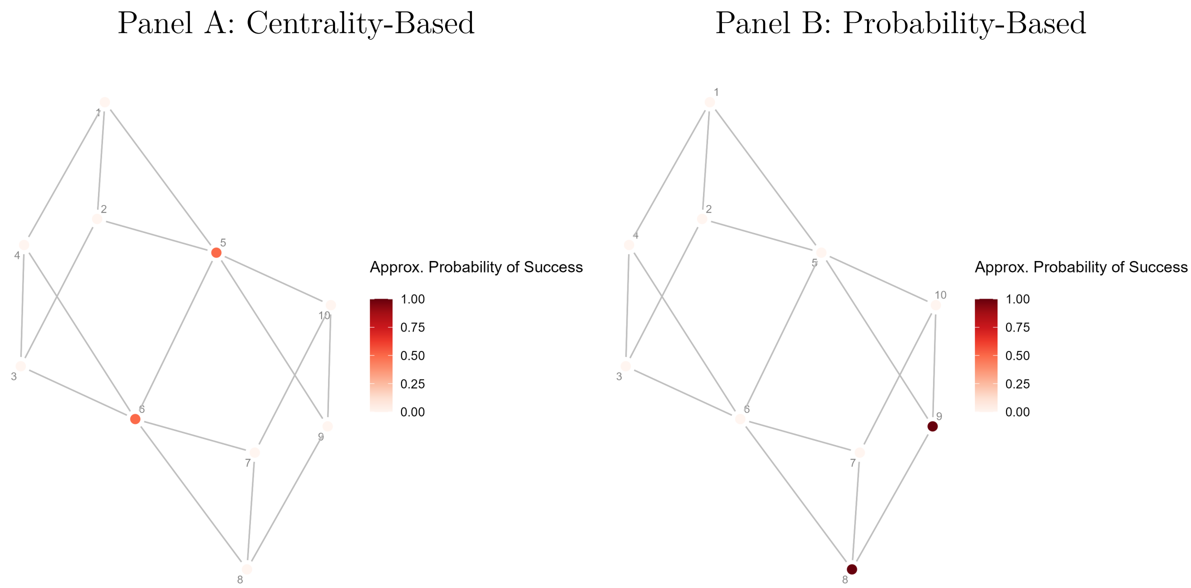


Figure 3: Initial Seeding based on Centrality and Probability

Panel B of Figure 3 captures seeding with probability-based targeting. The seeded households are the ones numbered 8 and 9. These households are selected as they have the highest $p^*_{iH}$s among all the households in this network. I can pick these households in simulations, as I can perfectly observe the households' $p^*_{iH}$s here. In practice, however, we may not have the necessary information to identify these households. In Appendix E, I discuss my strategy for estimating households' $p^*_{iH}$ 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.

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 convinced everyone else that satisfies equation (8) for this network to put effort into experimentation for getting informed. Using the efficiency measure defined in equation (9), I can score centrality-based targeting 0, while probability-based targeting scores 1. Therefore, the centrality-based targeting strategy fails in this scenario. It is also worth noting that in both types of targeting, in this scenario, the term $B_\kappa$ in (9) takes 0 as there is no mistargeting.

Panel A: Centrality-Based Panel B: Probability-Based

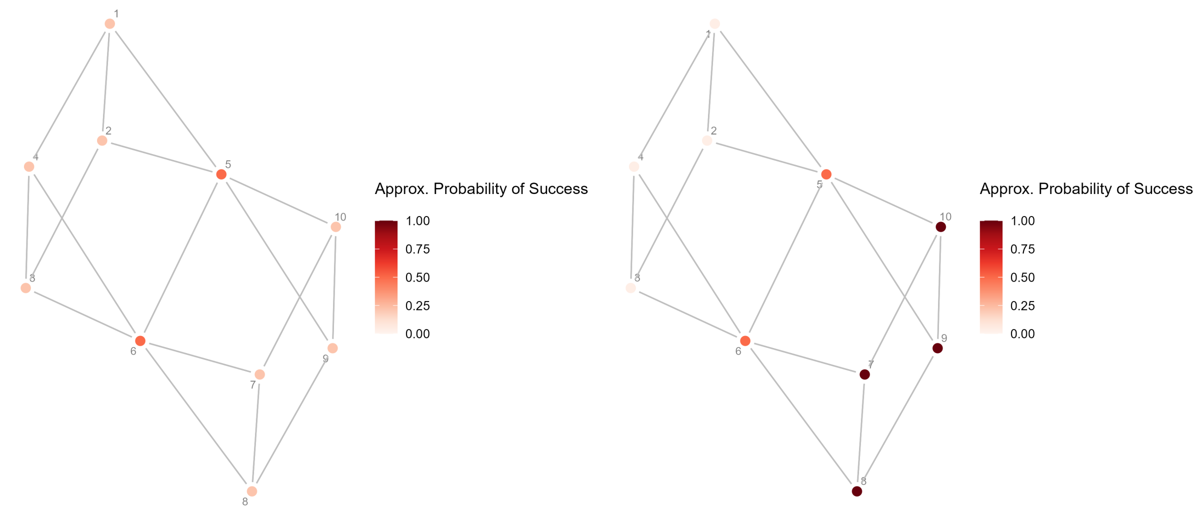


Figure 4: Performance of seeds after three periods

In this example, the $p^*_{iH}$s are highly heterogeneous, and the network connections are highly assortative in $p^*_{iH}$. In what follows, I first study non-assortative networks and understand the consequences of heterogeneity of $p^*_{iH}$s on the success of different targeting strategies. Then, I allow the networks to be perfectly assortative and vary the degree of heterogeneity in $p^*_{iH}$s. For these analyses, I simulate 200 networks with 30 households to assess the effectiveness of different targeting strategies on average.[25] As discussed above, the main focus is on centrality-based and probability-based targeting strategies. I also use a randomized targeting strategy (where seeds are selected randomly) for comparison.

[25]Appendix G presents the robustness of my results for networks with 20 and 40 households.

### 3.2 Targeting Homogeneous vs. Heterogeneous Networks

Let me focus on the consequences of heterogeneity in $p^*_{iH}$s for different network-based targeting strategies. Column (1) of Table 1 presents simulation results for networks with homogeneous $p^*_{iH} = p^*_H$. These networks are non-assortative in the probability.[26] Centrality-based targeting performs better than probability-based and random targeting for these networks.

I should note a few things in this regard. First, in networks with homogeneous $p^*_{iH} = p^*_H$, everyone should adopt or not adopt the technology given the same threshold probability of adoption. For the results presented here, I assume this threshold to be 0.4 for all households (Appendix G includes robustness of my results concerning change in this value). Thus, if $p^*_H \geq 0.4$, everyone should adopt the technology under efficient diffusion of information. Therefore, everyone optimally decides not to adopt the technology for a subsection of the simulated networks, where the randomly drawn $p^*_H < 0.4$. In terms of equation (9), these networks have $\mathit{Informed}^T = 0$. Similarly, for other simulated networks, where the randomly drawn $p^*_H \geq 0.4$, I have $\mathit{Uninformed}^T = 0$. As either $\mathit{Informed}^T$ or $\mathit{Uninformed}^T$ is zero for homogeneous networks, I cannot use $\mathit{Efficiency}_\kappa$ to measure the efficiency of the targeting strategy $\kappa$. Instead, I use the first term ($A_\kappa$) of $\mathit{Efficiency}_\kappa$ for that purpose, which leads to dropping networks with $p^*_H < 0.4$ from the average efficiency score calculation. To maintain parity with the number of observations in other columns, I increase the number of simulated networks to 400 for homogeneous networks. Finally, since everyone has the same $p^*_{iH} = p^*_H$ in a homogeneous network, the probability-based targeting reduces to systematically picking the first pair of households in the networks as the seeds.

In column (2) of Table 1, I allow the networks to be heterogeneous concerning $p^*_{iH}$s while remaining non-assortative. The efficiency scores become close to zero for all targeting strategies. The result aligns with my prediction in Figure 1. It is due to the diffusion being dependent on the $p^*_{iH}$s and the network ties not accounting for the heterogeneity in these probabilities. It is worth noting that the result is due to a lower value of $A_\kappa$ and a

[26] For a homogeneous network, assortivity property will lead to everyone being connected to everyone else.

Table 1: Efficiency Scores for Simulations using Different Targeting Strategies

| | | Homogeneous | Heterogeneous | |
|---|---|---|---|---|
| | | | Non-Assortative | Assortative |
| Targeting Strategy | Statistic | (1) | (2) | (3) |
| Eigenvector Centrality-Based | Mean | 0.455 | -0.003 | 0.412 |
| | Variance | 0.223 | 0.002 | 0.228 |
| Probability-Based | Mean | 0.189 | -0.040 | 0.956 |
| | Variance | 0.125 | 0.023 | 0.004 |
| Random | Mean | 0.000 | 0.000 | 0.438 |
| | Variance | 0.000 | 0.000 | 0.228 |
| | Observations† | 239 | 200 | 200 |

*Notes:*† Simulations are done for 400 networks with homogeneous probabilities and 200 networks with heterogeneous probabilities. Upon generation of the true probabilities, some networks are dropped as they contained 0% of informed households under full efficiency. Columns (2) and (3) use the efficiency measure $Efficiency_\kappa$ to measure the efficiency of the targeting strategy $\kappa$. Column (1) uses the term $A_\kappa$ of $Efficiency_\kappa$ for that purpose. All networks contain 30 households, and the threshold probability of learning is assumed to be 0.4 for all of them. For assortative networks, each pair of households having a success probability difference of 0.1 or less is assumed to be connected.

higher value of $B_\kappa$ for the measure in equation (9). Thus, targeting fails to reach agents that should adopt and mistargeting increases. Keeping the heterogeneity at the same level, I allow the networks to be assortative in the $p^*_{iH}$s for column (3). The effectiveness of all targeting strategies increases as a result. This result is because network ties are based on heterogeneity in $p^*_{iH}$s. As a result, we always reach agents with $p^*_{iH}$s similar to the initial seeds. Although seeds vary in effectiveness depending on their selection, all types perform better than they did for non-assortative networks.

More importantly, for column (3), probability-based seeds perform better than centrality-based seeds. Both of them perform better than the random seeds. By design, probability-based seeds target the population most likely to adopt due to the highest $p^*_{iH}$s. On the contrary, centrality-based seeds target the most influential agents in their connections. With a high level of heterogeneity in $p^*_{iH}$s (as for columns (2) and (3)), assortative networks imply that centrality-based seeds represent agents with average $p^*_{iH}$s. Thus, these agents are less effective than probability-based seeds in reaching the agents having the highest $p^*_{iH}$s. In the following sub-section, I study assortative networks with varying degrees of heterogeneity in $p^*_{iH}$s. The objective is to understand, for different network-based strategies, the role of such heterogeneity for perfectly assortative networks.

### 3.3 Targeting Assortative Networks with Varying Heterogeneity

I consider the agents to be connected for perfectly assortative networks if their $p^*_{iH}$s are within a difference of 0.1. Following the notation used in Section 2, this implies $\delta = 0.1$. I present the robustness of my results for different values of $\delta$ in Appendix G.[27] Figure 5 presents the performance of different targeting strategies with assortative networks over varying degrees of heterogeneity in $p^*_{iH}$s. Panel A of the figure presents the results in linear scale. Panel B shows the same results in a logarithmic scale for better visualization of efficiency scores with lower levels of heterogeneity.

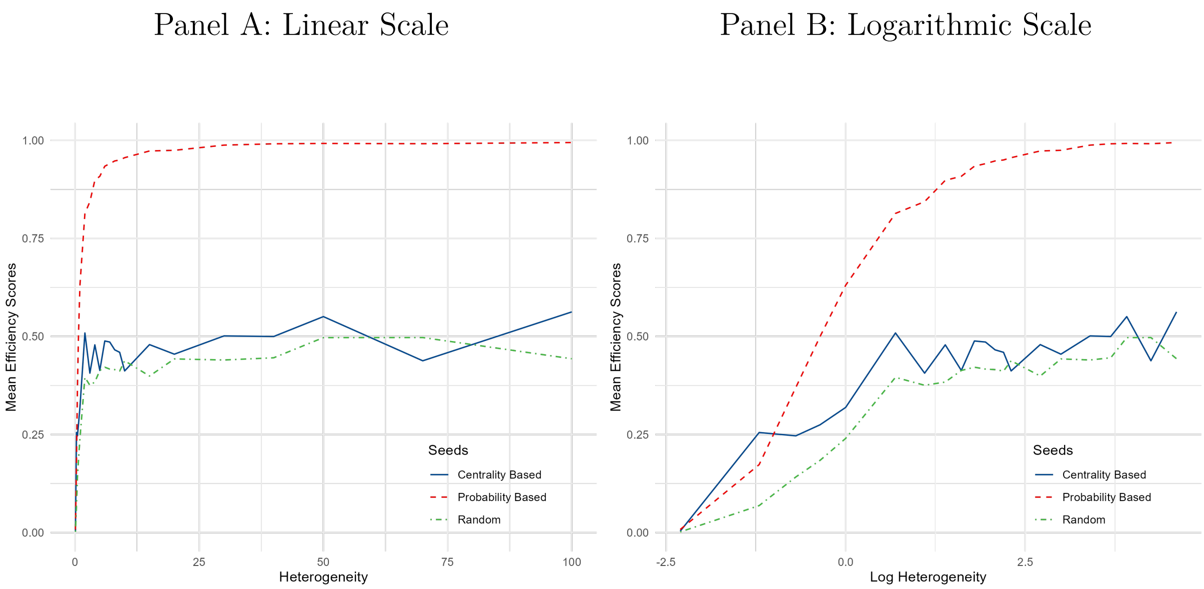


Figure 5: Efficiency scores over increasing levels of heterogeneity (with assortative networks)

As we can see, the performances of different targeting strategies improve with an increase in heterogeneity. As heterogeneity approximates to zero (i.e., almost converges to homogeneity), all targeting strategies approach an efficiency score of 0. The result is not surprising. For low levels of heterogeneity, everyone is connected in an assortative network. Thus, all three types of targeting reach everyone in the population, leading to a high value of both $A_\kappa$ and $B_\kappa$. Thus, the value of $Efficiency_\kappa$ converges to zero for all targeting strategies. As heterogeneity increases, we start observing the differences between the performances of different targeting strategies.

[27] I also show in the Figure G.8 of Appendix G that for intermediary levels of $\delta$, the results do not vary much for a given level of heterogeneity in $p^*_{iH}$s.

For lower levels of heterogeneity, centrality-based and probability-based targeting perform very similarly and slightly better than the random targeting strategy. Thus, for low heterogeneity in $p^*_{iH}$s, centrality-based targeting does not suffer (compared to other targeting strategies) for not accounting for this heterogeneity in its design. However, as heterogeneity increases, the negative effect of such heterogeneity on the performance of centrality-based targeting becomes apparent. As a result, probability-based targeting performs substantially better as heterogeneity increases than its centrality-based counterpart. Beyond a certain level of heterogeneity, the simulated networks converge to the maximum level of heterogeneity, and the average efficiency scores also converge to their maximum.

Let me now explain the results with the highest levels of heterogeneity. As the heterogeneity approaches its peak, everyone in the networks has a $p^*_{iH}$ of either 0 or 1. Given the assortative nature of these networks, everyone with $p^*_{iH} = 0$ is connected. Similarly, everyone with $p^*_{iH} = 1$ is connected. Depending on the number of households with 0 and 1 $p^*_{iH}$s, centrality-based targeting reaches either one of the groups. If more households have a $p^*_{iH}$ of 1, centrality-based targeting reaches all households with a $p^*_{iH}$ of 1. Conversely, centrality-based targeting reaches all households with a $p^*_{iH}$ of 0 if more households have a $p^*_{iH}$ of 0. Given the random nature of the simulations, either of these cases happens half of the time. Thus, centrality-based targeting converges to an efficiency score of around 50%.On the contrary, probability-based targeting always reaches households with $p^*_{iH} = 1$, leading to its convergence to an efficiency score of around 100%.

These results show that centrality-based targeting performs worse than probability-based targeting in reaching households with the highest $p^*_{iH}$ in networks assortative in these probabilities. However, the level of heterogeneity in $p^*_{iH}$ matters in this comparison. For low levels of this heterogeneity, both strategies perform similarly. The difference between them becomes prominent only when the heterogeneity increases beyond a certain threshold.

# 4 Empirical Strategy

My next objective is empirically testing the following hypotheses derived from the 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.*

Hypothesis 1 does not require networks to be assortative in the heterogeneity in benefits. It focuses on the underlying condition for the failure of centrality-based targeting in a complex diffusion process. As I show in Table 1, even with non-assortative networks, I expect the hypothesis to be true as long as the heterogeneity in benefits affects the diffusion process. On the other hand, Hypothesis 2 requires the assortativity property to be true. In my simulations, under the assumption of perfectly assortative networks, I show that probability-based seeds perform better than their centrality-based counterparts as heterogeneity increases. In reality, the networks are less likely to be perfectly assortative and more likely to be probabilistically assortative (where two agents with similar success probabilities are more *likely* to be connected, similar to the definition of *homophily* in Golub and Jackson (2012a,b,c)). Unfortunately, I do not observe network connections in the data I use in this study. As a result, I cannot assess these networks' degree of assortativity. However, accepting Hypothesis 2 would mean the existence of some degree of assortativity in the networks that we can use for policy purposes.

## 4.1 Data Sources

For my empirical analysis, I use the replication data from BBMM together with the survey data from AESTAS conducted by IFPRI. I briefly describe these datasets in this subsection before describing my identification strategies next.[28]

### 4.1.1 Replication data of BBMM

To promote *Pit Planting* (PP) for Maize farmers in Malawi, BBMM conducts a Randomized Controlled Trial (RCT) in 200 villages from 3 Malawian districts with semi-arid climates.

[28] A detailed description of these data sources are in Appendix C.

In each village, the researchers selected two *seed* farmers and trained them on PP. The criteria for the selection of these seed farmers were decided at the village level after the villages were randomly allocated into one of the four treatment arms:

1. **Complex Contagion Arm:** where the seed farmers are selected to maximize diffusion of information, assuming complex diffusion to be the underlying diffusion process.

2. **Simple Contagion Arm:** seed farmers are selected to maximize diffusion of information, assuming simple diffusion as the underlying diffusion process.[29]

3. **Geographic Arm:** seed farmers are selected to maximize the diffusion of information using geographic networks.

4. **Benchmark Arm:** seed farmers are selected by extension agents without using any network data.

BBMM uses simulations with baseline network information to identify two seeds optimal for improving the diffusion of information in all 200 villages, given an underlying diffusion process (complex contagion, simple contagion, and geographic). Once the villages were randomly allocated to one of these three treatment arms, they selected two households as seeds in a village, depending on the treatment arm allocated. For example, if the village got assigned to the *Complex Contagion Arm*, they selected as seeds the households identified in the simulation as the ones optimizing the diffusion of information in that village following a complex contagion diffusion process. If the village got allocated to the *Benchmark Arm*, the extension agents would select seed households instead. After training these seed farmers, they collected household survey data on farming techniques, input use, yields, assets, and other characteristics for a random panel of approximately 30 households per village (including the seed farmers). This led to the data on around 5600 households from the 200 villages for 2-3 survey rounds.

BBMM assumes the information of PP to be equally beneficial to all households in a village, with the heterogeneity in adoption being explained by differences in the adoption

[29] In a simple diffusion process, the information diffuses from one household to its connections with a random probability.

costs. In this study, I relax that assumption and consider heterogeneity in benefits and its consequences on the performance of their seeding strategies. They collected detailed data on household-level adoption decisions over multiple survey rounds, which I use to calculate the dependent variables for my analysis. Their replication data also includes information on household-level measures of centrality used to select seeds under different treatment arms, which I use to assess the centrality of seed households in their experiment. Additionally, my analysis requires the surveyed households' ex-ante probability of adoption. This information is not available in their replication data as they did not consider the benefits of adoption to be different across households. For this purpose, I turn to the AESTAS dataset.

#### 4.1.2 AESTAS data

AESTAS is a nationally representative panel household survey conducted by the International Food Policy Research Institute (IFPRI) to monitor Malawi's lead farmer (LF) program. [30] The data was collected in waves 1 in 2016 and 2 in 2018. The publicly available version of the survey dataset contains information from *household interviews*, *lead farmer interviews*, and *community interviews*. For this study, I use the data collected through household interviews, which collected data on technology adoption and awareness, exposure to different technologies, access to extension services, and socioeconomic and demographic characteristics. Around 3000 households were surveyed in wave 1, with 2880 re-surveyed in wave 2 (with the attrition rate being around 4%).

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 pre-determined agricultural technologies and practices.

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

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

## 4.2 Identification Strategy

The main empirical results of this study use the non-experimental village-level variations in the BBMM data. I additionally report the robustness of my results using their experimental variation in the Appendix F. 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 particular, given the selection of seeds in the BBMM experiment, I calculate the seeds' average centrality and probability of adoption and use them in the following regression:

$$\begin{aligned}\text{Outcome}_{vt} &= \beta_0 + \beta_1\text{Centrality}_v + \beta_2\text{Probability}_v + \beta_3\text{Heterogeneity}_v \\ &+ \beta_4\text{Centrality}_v \times \text{Heterogeneity}_v + \beta_5\text{Probability}_v \times \text{Heterogeneity}_v + \lambda X_v + \zeta_t + \epsilon_{vt}.\end{aligned} \tag{10}$$

$Outcome_{vt}$ denotes some adoption-related outcome for village $v$ at time $t$. Like BBMM, I focus on the outcomes in years 2 and 3, as they argued that the outcome variables start reflecting the diffusion of information from year 2 onwards. I defer the discussion on these outcome variables to the next section. $Centrality_v$ represents the average centrality of the seeds for village $v$, at the baseline. $Probability_v$ represents the average adoption probability for the seeds in village $v$ at the baseline, and $Heterogeneity_v$ is the baseline village-level heterogeneity in adoption probability. 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 ($\zeta_t$). The random error of the regression is captured by $\epsilon_{vt}$.

I calculate $Centrality_v$ using the seed households' eigenvector centrality at the baseline. The centrality measures are pre-calculated by BBMM and available in their replication data. However, the probability of adoption information is unavailable in the BBMM replication dataset as they do not consider households heterogeneous in the adoption benefits. I use the survey data from AESTAS to predict an adoption index conditional on demographics observed in both datasets.[31] Then, for the BBMM sample at the baseline,

[31]The details of this exercise are in Appendix E.

I calculate an out-of-sample prediction of this adoption index conditional on the same observable demographics.[32] I use this predicted adoption index as a proxy for the BBMM households' adoption probability.[33] $Probability_v$ is the average of this predicted adoption index for the seed households at the baseline. On the other hand, $Heterogeneity_v$ is the coefficient of variation of the same baseline predicted adoption index at the village level, capturing village-level heterogeneity in the predicted adoption index. 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.

The outcome variables (described in the next section), also used in the village-level analysis of BBMM, exclude the seeded households. I assume that the seed household characteristics (i.e., $Centrality_v$ and $Probability_v$ in regression (10)) are exogenous to the outcome variables. The assumption seems reasonable as the village-level outcomes do not consider the seeded households. Additionally, I assume that conditional on village level controls, $Heterogeneity_v$ is also exogenous in (10). However, as long as $Centrality_v$ and $Probability_v$ remain exogenous, no assumption is needed regarding the exogeneity of $Heterogeneity_v$ for identifying the coefficients of interest $\beta_4$ and $\beta_5$.

Endogenous $Centrality_v$ in (10) 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 (10). Similarly, the village-level unobserved characteristics affecting adoption-related outcomes should not be related to the

[32] A comparison of the BBMM and AESTAS samples are also in Appendix E.

[33] In the Appendices, I also provide the robustness of my results concerning a usage index (that uses the reported plot-level usage of technologies, instead of the said adoption of the same) as the proxy for households' adoption probability.

seed's adoption probability. As $Probability_v$ represents the average adoption probability of the seed households, it should also be exogenous in (10).[34] However, since I cannot verify these identifying assumptions, I check the robustness of my results using the experimental variations in the BBMM data that use weaker identifying assumptions. The Appendix F provides details.

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 others. In Appendix G, I check the robustness of my results by including the treatment dummies. As my results remain almost identical, I present them without the treatment dummies in the following section.

# 5 Results and Discussion

## 5.1 Descriptive Statistics

Table 2 describes 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 (10). The first four columns of the table represent the within-treatment group variations. Regression specification (12) in Appendix F uses the experimental variations between these four groups.

The first two rows present the outcome variables of my analysis, which are the same ones used in the village-level analysis of BBMM. 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.[35] Any Non-Seed Adopters (PP) is a dummy variable that captures whether

[34] 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 (10).

[35] Shadow farmers are seed farmers chosen by the BBMM simulation, assuming some underlying diffusion model. However, they were not selected as seeds in the BBMM experiment because their villages were assigned to receive seeding based on a different diffusion model.

Table 2: 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) |
| 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) |
| 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.

the villages had at least one *non-seed* farmer adopting pit planting in an agricultural season. 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 following two rows of table 2 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.[36] By the design of

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

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 (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 from the benchmark seeds. However, complex and geo seeds are statically different from the benchmark (at 1% and 5% levels of significance, respectively).

I use the predicted adoption index as the proxy for the adoption probability. Complex and benchmark seeds have the highest adoption probabilities, followed by simple seeds. The geo seeds have the lowest baseline probability of adoption. No statistically significant differences exist between benchmark, complex, and simple seeds. However, geo seeds are statistically different from their benchmark counterparts (at a 1% significance level).

The final row of table 2 presents 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 6. 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) with higher than the median average eigenvector centrality at baseline. Similarly, probability-based seed households (s) are defined to have higher than the median 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.[37] 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.[38]

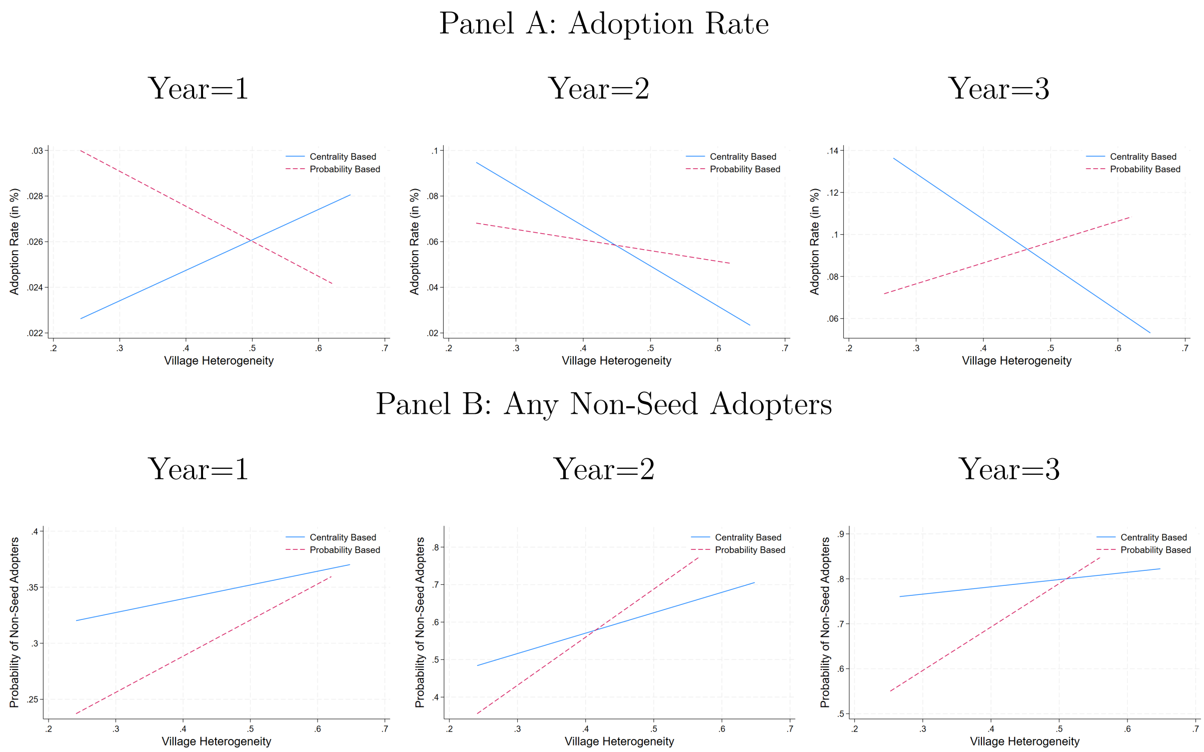


Figure 6: Outcomes for Different Seeding Strategies with respect to Village Heterogeneity

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

[37]Average village level correlation between the households' centrality and adoption probability, calculated at the baseline, is around 0.3. This is robust to using different centrality and adoption probability measures. Thus, a centrality-based seeding strategy should lead to a mostly different set of seed households than a probability-based strategy.

[38]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.

counterpart. The opposite is true for probability-based seeds compared to 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. Although, the centrality-based seeds remain the more successful for all levels of village heterogeneity.

Although informative, the descriptive figures do not account for village-level heterogeneity in other variables. In defining the centrality-based and probability-based seeds as dummy variables, Figure 6 also fails to capture the entire village-level variations of these seeds regarding their centrality and probability measures. In the following sub-section, I present the main results of my analysis that test my hypotheses more formally.

## 5.2 Main Results

Table 3 presents the main results of my analysis. Here, the main coefficients of interest are those corresponding to the interactions of $Heterogeneity_v$, with $Centrality_v$ and $Probability_v$. Following my hypotheses, 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 one standard deviation increase in eigenvector centrality for a completely homogeneous village is associated with a 1.47-1.88 standard deviation increase in the adoption rate. However, for villages with heterogeneity at the level of baseline mean, the effect drops to an increase of only 0.18-0.29 standard deviations. Similarly, one standard deviation increase in predicted adoption decreases the adoption rate by 1.28-1.78 standard deviations for a homogeneous village. But for villages with heterogeneity at the level of baseline mean, the effect drops to a decrease of 0.2-0.3 standard deviations only.

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 one standard deviation increase in eigenvector centrality for completely homogeneous villages leads to a 0.24-0.25 standard deviation increase in the probability of having at least one non-seed adopter. However, for villages with heterogeneity at the level of baseline mean, the effect drops to 0.02 standard deviations increase. On the other hand, one standard deviation increase in the predicted adoption index is associated with a 0.25-0.61 standard deviation decrease in the probability of having at least one non-seed adopter for a homogeneous village. For villages with heterogeneity at the level of baseline mean, however, the effect drops to a decrease of 0.04-0.08 standard deviations.

Table 3: 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* | 18.484*** | 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.

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 their education, skills, and ability, which affect how much they can learn about and use the new technology 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.

Following my simulation results, 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 increases, but targeting based on the adoption probability to perform better if the network is highly assortative in terms of the benefits from the technology. To test my 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. My 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 decreases 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 (reported in the Appendix F).

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, a randomized controlled trial that randomly allocates regions into a centrality-based, probability-based, and random seeding strategy would be more suitable. A randomized controlled trial of such a nature could also help in disentangling the effects of centrality and probability of seeds. These and 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.

Corazzini, L., F. Pavesi, B. Petrovich, and L. Stanca (2012): "Influential listeners: An experiment on persuasion bias in social networks," *European Economic Review*, 56, 1276–1288.

Crane-Droesch, A. (2017): "Technology Diffusion, Outcome Variability, and Social Learning: Evidence from a Field Experiment in Kenya," *American Journal of Agricultural Economics*, 100, 955–974.

de Janvry, A., M. Rao, and E. Sadoulet (2022): "Seeding the Seeds: Role of Social Structure in Agricultural Technology Diffusion," *Unpublished Manuscript.*

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.

——— (2012a): "Does Homophily Predict Consensus Times? Testing a Model of Network Structure via a Dynamic Process," *Review of Network Economics*, 11.

——— (2012b): "How Homophily Affects the Speed of Learning and Best-Response Dynamics," *The Quarterly Journal of Economics*, 127, 1287–1338.

——— (2012c): "Network Structure and the Speed of Learning Measuring Homophily Based on its Consequences," *Annals of Economics and Statistics*, 33.

Golub, B. and E. Sadler (2016): "Learning in Social Networks," in *The Oxford Handbook of the Economics of Networks*, ed. by Y. Bramoullé, A. Galeotti, and B. W. Rogers, Oxford University Press, 503–542.

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 $j$. [39] An *edge* $(i, j)$ is defined as a link from $i$ to $j$.[40] 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 $j$.[41]

**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

[39] 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.

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

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

the number of nodes the node $k$ connects to:

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

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:

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

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 $\lambda$ 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 $\lambda^R$ 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

For the simulations presented in this paper, I simulated three different categories of networks. These are homogeneous non-assortative networks, heterogeneous non-assortative networks, and heterogeneous assortative networks. In this section, I describe the methodology used for these simulations.

## B.1 Simulating Homogeneous Networks

**Step 1: Generating networks of households**

The first step in simulating homogeneous non-assortative undirected networks is to randomly generate symmetric adjacency matrix $g$ with elements 0 or 1 such that $g_{ij} = g_{ji}$, and $g_{ii} = 0$, $\forall i, j \in \mathcal{I}$. Then I generate the influence matrix $G$ by normalizing each row of g.[42] 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 networks.

**Step 2: Generating true probabilities of success**

The next step is to generate $p^*_{iH}$s for the networks. For homogeneous networks $p^*_{iH} = p^*_H$, $\forall i \in \mathcal{I}$. I draw one value of $p^*_H$ for each of the 200 networks from the uniform distribution $U(0, 1)$.

**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 we select in a given village? I consider three different targeting strategies:

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

For the homogeneous networks, the probability-based strategy will systematically select the first two households in a network, as all pairs of households have the same average $p^*_{iH}$s.

**Step 4: Diffusion**

Given the seeding strategy in a network, I let the diffusion take place for 10 periods. In

[42] Following convention, I assume the diagonal elements of the adjacency matrix ($g_{ii}$) to be zero. However, I set $g_{ii} = 1$ before calculating the influence matrix $G$ to allow for $G_{ii} \neq 0$ (for all networks in my simulation exercise, irrespective of whether they are homogeneous or heterogeneous). For calculating centrality measures (description below), the adjacency matrix $g$ with $g_{ii} = 0$ is used.

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 $\bar{p}^H_i = \bar{p}^H = 0.4$, 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**

In a set of 200 networks, I evaluate the targeting efficiency on average. Targeting efficiency of strategy $\kappa$ is measured by the following equation in each network:

$$\text{Efficiency}_\kappa = \underbrace{\frac{\text{Informed}^T_\kappa}{\text{Informed}^T}}_{A_\kappa} - \underbrace{\frac{\text{Informed}^F_\kappa}{\text{Uninformed}^T}}_{B_\kappa}$$

Here Informed$^T$ denotes the number of non-seed households that should get informed as they would adopt the technology under perfect information. Additionally, Informed$^T_\kappa$ captures the number of non-seed households that get informed within 10 periods of implementing the targeting strategy $\kappa$, among those households in Informed$^T$. Uninformed$^T$ denotes the number of non-seed households that should not get informed, and Informed$^F_\kappa$ captures the number of non-seed households that end up getting informed among those households given targeting strategy $\kappa$.

**Step 6: Comparison**

The evaluation is done for different seeding strategies. The results are then compared.

## B.2 Simulating Heterogeneous Non-Assortative Networks

For heterogeneous non-assortative networks, step 1 and steps 3-6 remain the same. The only difference is in step 2, where $p^*_{iH}$s (different for each households in heterogeneous networks) are drawn for each household $i \in \mathcal{I}$ independently from the normal distribution $N(0.5, 10)$. The draws are then truncated such that $p^*_{iH} \geq 1$ is truncated to 1, and $p^*_{iH} \leq 0$ is truncated to 0.

## B.3 Simulating Heterogeneous Assortative Networks

Compared to homogeneous non-assortative networks, heterogeneous assortative networks differ in steps 1 and 2. The rest of the steps remain the same.

**Step 1: Generating true probabilities of success**

For heterogeneous assortative networks, the first step is to generate $p^*_{iH}$s for a network. For that purpose, I randomly draw $p^*_{iH}$s for each household in a network from the normal distribution $N(0.5, \sigma)$. The draws are then truncated, if necessary, such that $p^*_{iH} \geq 1$ is truncated to 1, and $p^*_{iH} \leq 0$ is truncated to 0. If I set $\sigma$ to be large enough, it would lead all $p^*_{iH}$s to be either 0 or 1. On the other hand, lower values of $\sigma$ keep $p^*_{iH}$s more within 0 and 1. So, I can vary $\sigma$ to control the degree of heterogeneity in terms of $p^*_{iH}$s. This procedure is repeated 200 times for each $\sigma$ to generate 200 villages with differing levels of heterogeneity in terms of $p^*_{iH}$s, independent from each other. For Table 1 and its robustness checks $\sigma = 10$, to make it comparable with the heterogeneous non-assortative networks whose $p^*_{iH}$s are drawn from the normal distribution $N(0.5, 10)$. For Figure 5 and its robustness checks, $\sigma$ takes a wider range of values within 0.1 and 100 (both inclusive).

**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 $\forall i \neq j$, $g_{ij} = 1$ if $|p^*_{iH} - p^*_{jH}| < 0.1$ and 0 otherwise, and $g_{ii} = 0$. I then generate the influence matrix $G$ by normalizing each row of g (following the same methodology described for the homogeneous networks above). This procedure is repeated to generate 200 village networks for each value of $\sigma$.

# C Detailed Description of the Data Sources

## C.1 Replication data of BBMM

In the BBMM experiment, the researchers conducted a Randomized Controlled Trial (RCT) to promote *Pit Planting* (PP) for Maize farmers in Malawi.[43] 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.[44]

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.

[43]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.

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

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 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. 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.[45] For more details on the intervention and sampling of the study, please consult BBMM.

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.

## C.2 AESTAS data

AESTAS covered all 29 districts of Malawi, except Likoma.[46] The data collection was done in waves 1 in 2016 and 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.[47] Stratification was done based on whether or not the household had a LF. Per section, up to two households

[45] 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.

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

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

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.

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

D for details on the construction of these indices). I use these indices as proxies for the probability of adoption.

# D 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.

# E 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 $\text{Probability}_v$ in the regressions, while the coefficient of variation of this measure at the village level is $\text{Heterogeneity}_v$. However, in the BBMM experiment, the researchers did not collect any information about these probabilities, as their micro-foundations assumed the new technology's benefits were 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 D 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:

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

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 subsection to other household characteristics observable only in the AESTAS data and not in the BBMM data. The term $\mu_{it}$ 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 squares). However, I check the robustness of my results to non-linear specifications. I present these in Appendix G. 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.[48] 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 (12), as the identification uses experimental variations. However, we must consider the consequences for (10). The bias in estimating households' adoption probabilities would lead to a biased $\text{Probability}_v$ in (10). However, this will only create a problem in identifying the coefficient of interest $\beta_5$ 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.

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

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 (12) provides an alternative.

Table E.1: 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 E.1. The comparison is important as it helps me understand how the estimates 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 focuses only on the Maize farmers.

Table E.2: 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*** | 0.010*** | 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 E.2 presents the main results for this subsection. Here, I estimate the adoption and usage indices conditional on the demographics presented in Table E.1. 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.[49] 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. The magnitudes and significance levels vary, controlling for other household-level characteristics.

[49]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.

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.

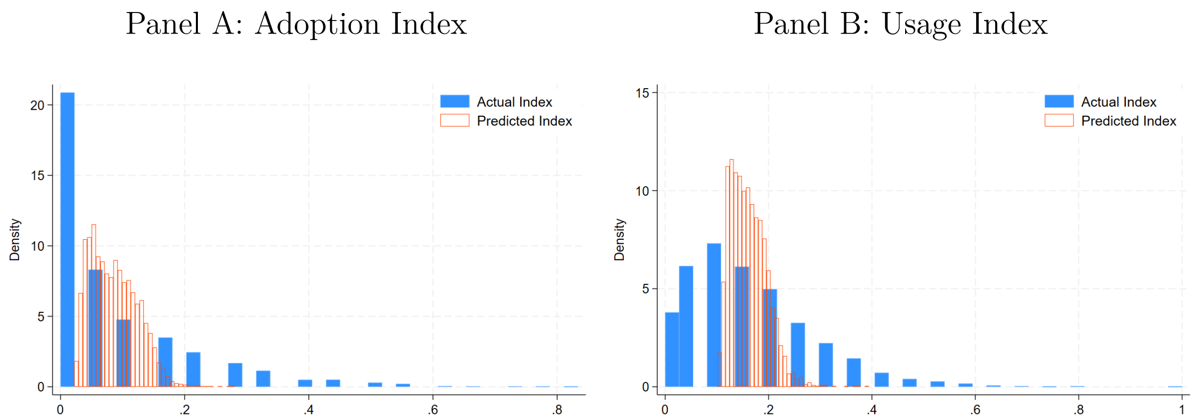


Figure E.1: Actual and Predicted Adoption and Usage Indices

Figure E.1 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. 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.

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 for unbiased identification of $\beta_5$ in (10), as I already discussed in the last section.

# F Identification using Between-Treatment Variations

## F.1 Specification

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

$$\begin{aligned}
\text{Outcome}_{vt} &= \theta_0 + \theta_1 \text{Centrality}_v + \theta_2 \text{Probability}_v + \theta_3 \text{Heterogeneity}_v \\
&+ \xi_b \text{Centrality}_v \times \text{Heterogeneity}_v + \xi_c \text{Centrality}_v \times \text{Heterogeneity}_v \times \text{Complex}_v \\
&+ \xi_s \text{Centrality}_v \times \text{Heterogeneity}_v \times \text{Simple}_v + \xi_g \text{Centrality}_v \times \text{Heterogeneity}_v \times \text{Geo}_v \\
&+ \phi_b \text{Probability}_v \times \text{Heterogeneity}_v + \phi_c \text{Probability}_v \times \text{Heterogeneity}_v \times \text{Complex}_v \\
&+ \phi_s \text{Probability}_v \times \text{Heterogeneity}_v \times \text{Simple}_v + \phi_g \text{Probability}_v \times \text{Heterogeneity}_v \times \text{Geo}_v \\
&+ \gamma X_v + \rho_t + \eta_{vt}.
\end{aligned} \tag{12}$$

Specification (12) is similar to specification (10), except the interactions of Centrality$_v$ $\times$ Heterogeneity$_v$ and Probability$_v$ $\times$ Heterogeneity$_v$ with treatment dummies. Here, $\xi_b$ captures the interaction between seed centrality and village-level heterogeneity for the benchmark treatment group. $\xi_c$, $\xi_s$, and $\xi_g$ captures how that interaction changes compared to the benchmark for complex, simple, and geo treatment groups. Similarly, $\phi_b$ captures the interaction between seed probability and village-level heterogeneity for the benchmark treatment group. $\phi_c$, $\phi_s$, and $\phi_g$ captures how that interaction changes 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, I am more interested in exploring between-group variations using this specification. Thus, the 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 with the same heterogeneity level 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 diffuses the technology to fewer (more) households. While exploring the village-level non-experimental variations, I argue the same for (10). 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 $\xi_c$ to be positive (negative) if the complex treatment group has more (less) central seeds than the benchmark. Similarly, I expect $\phi_c$ 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 $\xi_c$; 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 $\phi_c$; 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 $\xi_c$; 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 $\phi_c$; seeds with higher adoption probability than benchmark: depends on the relative effects of the increase in probability and heterogeneity.

Like specification (10), 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 (12) as BBMM argues that the treatment status affects the outcome variables only through the centrality of the seeds. In Appendix G, I present the robustness of my results by including the treatment dummies. My results remain robust.

## F.2 Results

Table F.1 focuses on exploring between treatment group variations. Here, I am interested in the coefficients of $\text{Centrality}_v \times \text{Heterogeneity}_v$ and $\text{Probability}_v \times \text{Heterogeneity}_v$, across different treatment groups. Note that the sign of $\text{Centrality}_v \times \text{Heterogeneity}_v$ is negative, and the sign of $\text{Probability}_v \times \text{Heterogeneity}_v$ is positive within different treatment groups, in line with the results of table 3. Table F.1 notes the differences in the coefficients of $\text{Centrality}_v \times \text{Heterogeneity}_v$ and $\text{Probability}_v \times \text{Heterogeneity}_v$, 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 subsection.

Columns (1) and (2) 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 eigenvector centrality of seed households leads to a 1.01-1.24 standard deviation improvement in the adoption rate. However, for benchmark villages having heterogeneity at the level of baseline benchmark mean, the effect drops to a decrease of 0.27-0.38 standard deviations. 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 0.95-1.42 standard deviations for a homogeneous village. However, for benchmark villages having

Table F.1: Village level Regression 2 of Adoption Outcomes (Pit Planting)

| Variables | Adoption Rate (1) | (2) | Any Non-Seed Adopters (3) | (4) |
|---|---|---|---|---|
| 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 | -10.419*** | -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$_v$ × Heterogeneity$_v$ | -2.423** | -2.237** | -6.692 | -6.574 |
| | (1.093) | (0.996) | (4.503) | (4.119) |
| Centrality$_v$ × Heterogeneity$_v$ × Complex | 0.657** | 0.664** | 4.328** | 3.756** |
| | (0.306) | (0.282) | (1.775) | (1.664) |
| Centrality$_v$ × Heterogeneity$_v$ × Simple | 0.416 | 0.428 | 1.078 | 0.431 |
| | (0.337) | (0.320) | (2.060) | (1.947) |
| Centrality$_v$ × Heterogeneity$_v$ × Geo | 2.026** | 1.942** | 0.103 | -0.070 |
| | (0.940) | (0.839) | (2.235) | (2.098) |
| Probability$_v$ × Heterogeneity$_v$ | 5.881** | 4.104* | 22.97*** | 12.35 |
| | (2.437) | (2.286) | (7.720) | (7.626) |
| Probability$_v$ × Heterogeneity$_v$ × Complex | -0.155 | -0.232 | -1.275 | -0.679 |
| | (0.520) | (0.497) | (2.765) | (2.654) |
| Probability$_v$ × Heterogeneity$_v$ × Simple | -0.121 | -0.110 | 1.941 | 3.511 |
| | (0.642) | (0.571) | (3.572) | (3.333) |
| Probability$_v$ × Heterogeneity$_v$ × 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.

heterogeneity at the baseline benchmark mean level, the effect size decreases to 0-0.04 standard deviations. The positive effect 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. The impact is significantly lower compared to the benchmark for the geo treatment group only.

The results for Any Non-Seed Adopters are in columns (3) and (4), with and without the village-level controls. For this outcome variable, the effects are also in the same direction for all the treatment groups. The results show that one standard deviation increase in eigenvector centrality for homogeneous villages leads to a 0.33-0.34 standard deviation improvement in the probability of having at least one non-seed adopter. But, for benchmark villages having heterogeneity at the level of baseline benchmark mean, the effect drops to around 0.18 standard deviations decrease in the probability of having at least one non-seed adopter. The negative impact of heterogeneity on the relationship between seeds' centrality and the probability of having non-seed adopters is lower for the other treatment groups. However, the effect is significantly lower only for the complex treatment group. 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 0.45-0.79 standard deviations for a homogeneous village. However, for benchmark villages having heterogeneity at the level of baseline benchmark mean, the effect drops to a decrease of 0.09-0.11 standard deviations. 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.

# G Robustness Checks

Table G.1: Simulation Robustness (w.r.t different centrality measure)

| | | Homogeneous | Heterogeneous | |
|---|---|---|---|---|
| | | | Non-Assortative | Assortative |
| Targeting Strategy | Statistic | (1) | (2) | (3) |
| Betweenness Centrality-Based | Mean | 0.463 | -0.010 | 0.635 |
| | Variance | 0.225 | 0.002 | 0.210 |
| Probability-Based | Mean | 0.189 | -0.040 | 0.956 |
| | Variance | 0.125 | 0.023 | 0.004 |
| Random | Mean | 0.000 | 0.000 | 0.438 |
| | Variance | 0.000 | 0.000 | 0.228 |
| | Observations$^{\dagger}$ | 239 | 200 | 200 |

*Notes:*$^{\dagger}$ Simulations are done for 400 networks with homogeneous probabilities and 200 networks with heterogeneous probabilities. Upon generation of the true probabilities, some networks are dropped as they contained 0% of informed households under full efficiency. Columns (2) and (3) use the efficiency measure $\text{Efficiency}_{\kappa}$ to measure the efficiency of the targeting strategy $\kappa$. Column (1) uses the term $A_{\kappa}$ of $\text{Efficiency}_{\kappa}$ for that purpose. All networks contain 30 households, and the threshold probability of learning is assumed to be 0.4 for all of them. For assortative networks, each pair of households having a success probability difference of 0.1 or less is assumed to be connected.

Table G.2: Simulation Robustness (w.r.t $\bar{p}_i^H = 0.5$, instead of $\bar{p}_i^H = 0.4$)

| | | Homogeneous | Heterogeneous | |
|---|---|---|---|---|
| | | | Non-Assortative | Assortative |
| Targeting Strategy | Statistic | (1) | (2) | (3) |
| Eigenvector Centrality-Based | Mean | 0.197 | -0.007 | 0.414 |
| | Variance | 0.136 | 0.006 | 0.230 |
| Probability-Based | Mean | 0.017 | -0.009 | 0.965 |
| | Variance | 0.008 | 0.012 | 0.003 |
| Random | Mean | 0.000 | 0.000 | 0.161 |
| | Variance | 0.000 | 0.000 | 0.129 |
| | Observations$^{\dagger}$ | 197 | 200 | 200 |

*Notes:*$^{\dagger}$ Simulations are done for 400 networks with homogeneous probabilities and 200 networks with heterogeneous probabilities. Upon generation of the true probabilities, some networks are dropped as they contained 0% of informed households under full efficiency. Columns (2) and (3) use the efficiency measure Efficiency$_{\kappa}$ to measure the efficiency of the targeting strategy $\kappa$. Column (1) uses the term $A_{\kappa}$ of Efficiency$_{\kappa}$ for that purpose. All networks contain 30 households, and the threshold probability of learning is assumed to be 0.5 for all of them. For assortative networks, each pair of households having a success probability difference of 0.1 or less is assumed to be connected.

Table G.3: Simulation Robustness (w.r.t $\bar{p}_i^H = 0.3$, instead of $\bar{p}_i^H = 0.4$)

| | | Homogeneous | Heterogeneous | |
|---|---|---|---|---|
| | | | Non-Assortative | Assortative |
| Targeting Strategy | Statistic | (1) | (2) | (3) |
| Eigenvector Centrality-Based | Mean | 0.642 | -0.004 | 0.409 |
| | Variance | 0.218 | 0.008 | 0.224 |
| Probability-Based | Mean | 0.481 | -0.031 | 0.948 |
| | Variance | 0.236 | 0.012 | 0.004 |
| Random | Mean | 0.018 | 0.003 | 0.469 |
| | Variance | 0.010 | 0.003 | 0.227 |
| | Observations$^{\dagger}$ | 281 | 200 | 200 |

*Notes:*$^{\dagger}$ Simulations are done for 400 networks with homogeneous probabilities and 200 networks with heterogeneous probabilities. Upon generation of the true probabilities, some networks are dropped as they contained 0% of informed households under full efficiency. Columns (2) and (3) use the efficiency measure Efficiency$_{\kappa}$ to measure the efficiency of the targeting strategy $\kappa$. Column (1) uses the term $A_{\kappa}$ of Efficiency$_{\kappa}$ for that purpose. All networks contain 30 households, and the threshold probability of learning is assumed to be 0.3 for all of them. For assortative networks, each pair of households having a success probability difference of 0.1 or less is assumed to be connected.

Table G.4: Simulation Robustness (w.r.t 20 households, instead of 30, per network)

| | | Homogeneous | Heterogeneous | |
|---|---|---|---|---|
| | | | Non-Assortative | Assortative |
| Targeting Strategy | Statistic | (1) | (2) | (3) |
| Eigenvector Centrality-Based | Mean | 0.724 | -0.029 | 0.464 |
| | Variance | 0.184 | 0.018 | 0.236 |
| Probability-Based | Mean | 0.504 | -0.072 | 0.947 |
| | Variance | 0.226 | 0.031 | 0.008 |
| Random | Mean | 0.025 | -0.014 | 0.447 |
| | Variance | 0.015 | 0.012 | 0.233 |
| | Observations† | 230 | 200 | 200 |

*Notes:*† Simulations are done for 400 networks with homogeneous probabilities and 200 networks with heterogeneous probabilities. Upon generation of the true probabilities, some networks are dropped as they contained 0% of informed households under full efficiency. Columns (2) and (3) use the efficiency measure $\text{Efficiency}_\kappa$ to measure the efficiency of the targeting strategy $\kappa$. Column (1) uses the term $A_\kappa$ of $\text{Efficiency}_\kappa$ for that purpose. All networks contain 20 households, and the threshold probability of learning is assumed to be 0.4 for all of them. For assortative networks, each pair of households having a success probability difference of 0.1 or less is assumed to be connected.

Table G.5: Simulation Robustness (w.r.t 40 households, instead of 30, per network)

| | | Homogeneous | Heterogeneous | |
|---|---|---|---|---|
| | | | Non-Assortative | Assortative |
| Targeting Strategy | Statistic | (1) | (2) | (3) |
| Eigenvector Centrality-Based | Mean | 0.184 | -0.002 | 0.504 |
| | Variance | 0.125 | 0.002 | 0.232 |
| Probability-Based | Mean | 0.013 | -0.009 | 0.955 |
| | Variance | 0.009 | 0.009 | 0.003 |
| Random | Mean | 0.000 | 0.000 | 0.103 |
| | Variance | 0.000 | 0.000 | 0.086 |
| | Observations† | 241 | 200 | 200 |

*Notes:*† Simulations are done for 400 networks with homogeneous probabilities and 200 networks with heterogeneous probabilities. Upon generation of the true probabilities, some networks are dropped as they contained 0% of informed households under full efficiency. Columns (2) and (3) use the efficiency measure $\text{Efficiency}_\kappa$ to measure the efficiency of the targeting strategy $\kappa$. Column (1) uses the term $A_\kappa$ of $\text{Efficiency}_\kappa$ for that purpose. All networks contain 40 households, and the threshold probability of learning is assumed to be 0.4 for all of them. For assortative networks, each pair of households having a success probability difference of 0.1 or less is assumed to be connected.

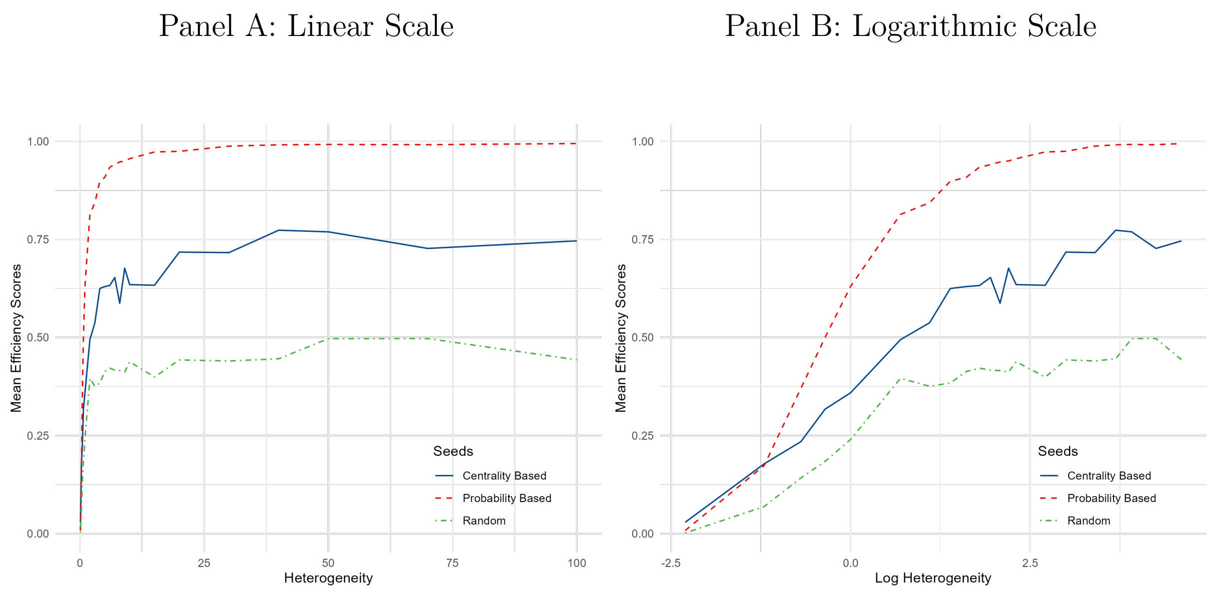


Figure G.1: Efficiency scores over increasing levels of heterogeneity (with assortative networks) w.r.t betweenness centrality (instead of eigenvector centrality)

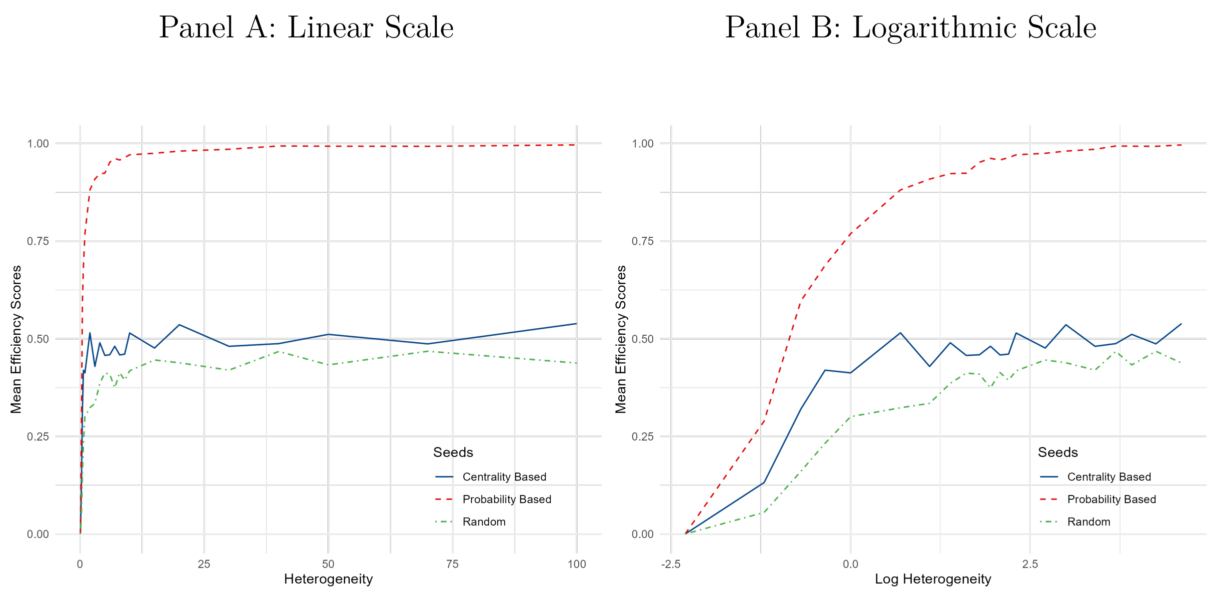


Figure G.2: Efficiency scores over increasing levels of heterogeneity (with assortative networks w.r.t $\delta = 0.2$ instead of $\delta = 0.1$)

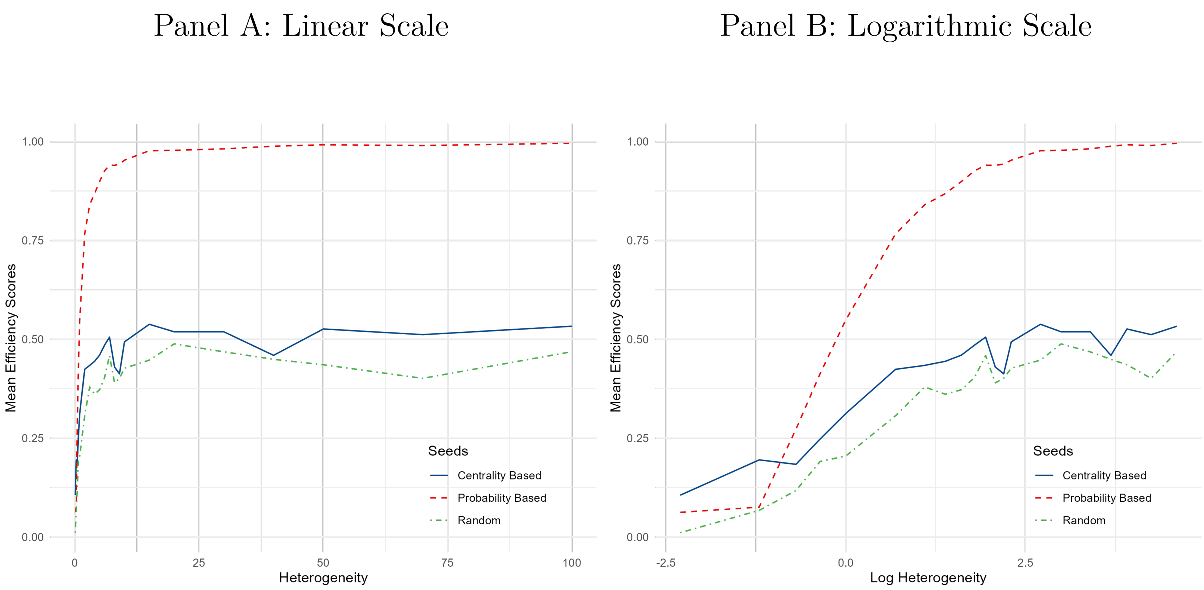


Figure G.3: Efficiency scores over increasing levels of heterogeneity (with assortative networks w.r.t $\delta = 0.05$ instead of $\delta = 0.1$)

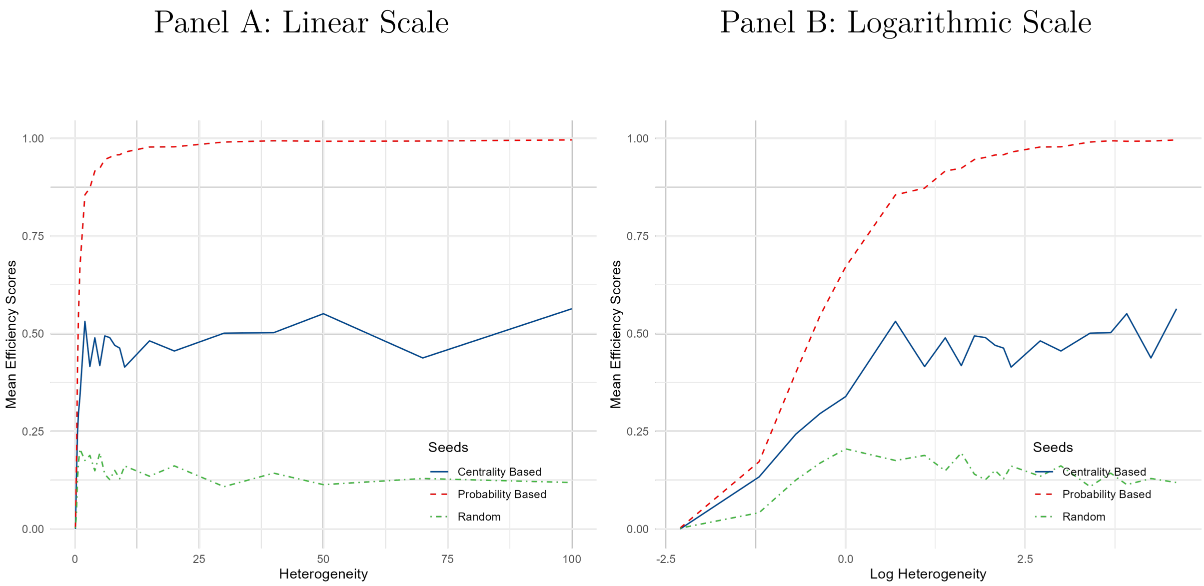


Figure G.4: Efficiency scores over increasing levels of heterogeneity (w.r.t $\bar{p}_i^H = 0.5$, instead of $\bar{p}_i^H = 0.4$)

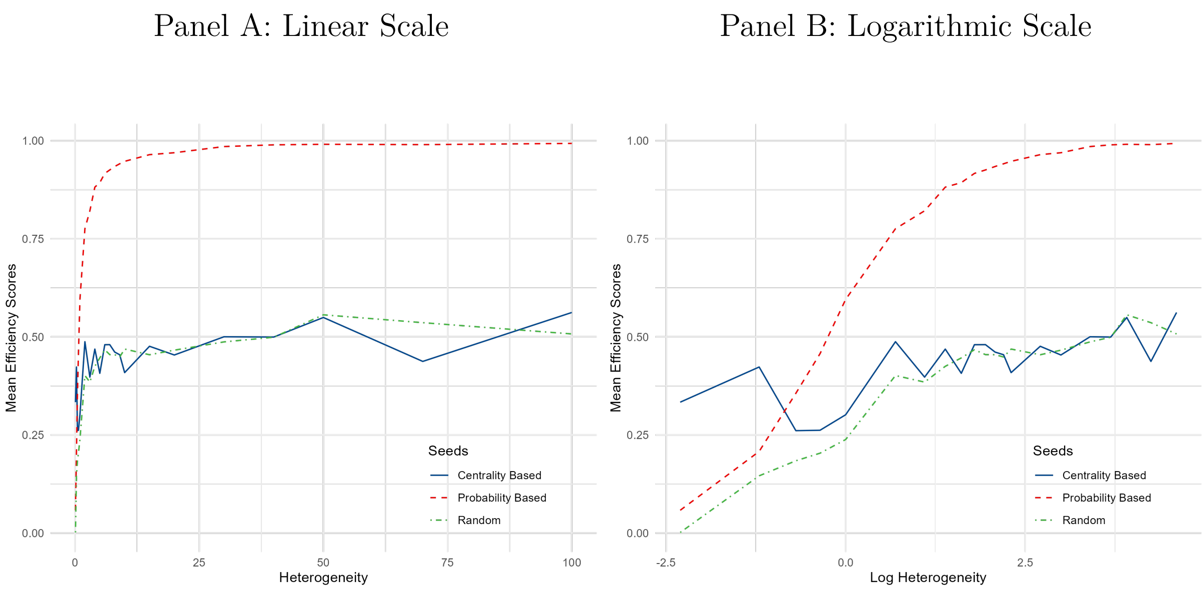


Figure G.5: Efficiency scores over increasing levels of heterogeneity (w.r.t $\bar{p}_i^H = 0.3$, instead of $\bar{p}_i^H = 0.4$)

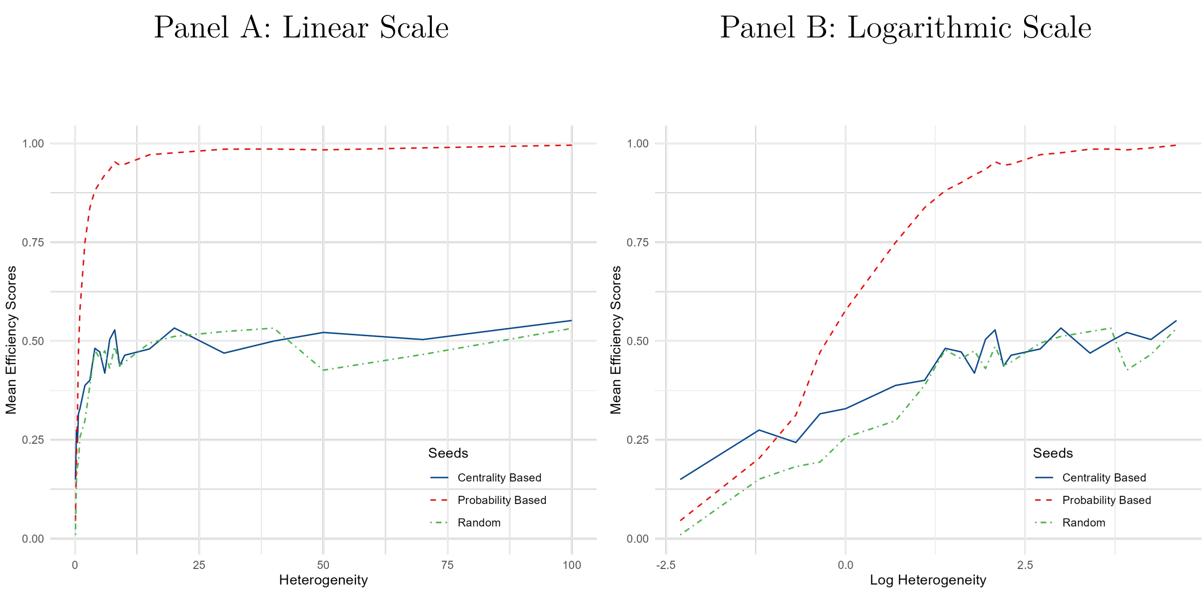


Figure G.6: Efficiency scores over increasing levels of heterogeneity (w.r.t 20 households, instead of 30, per network)

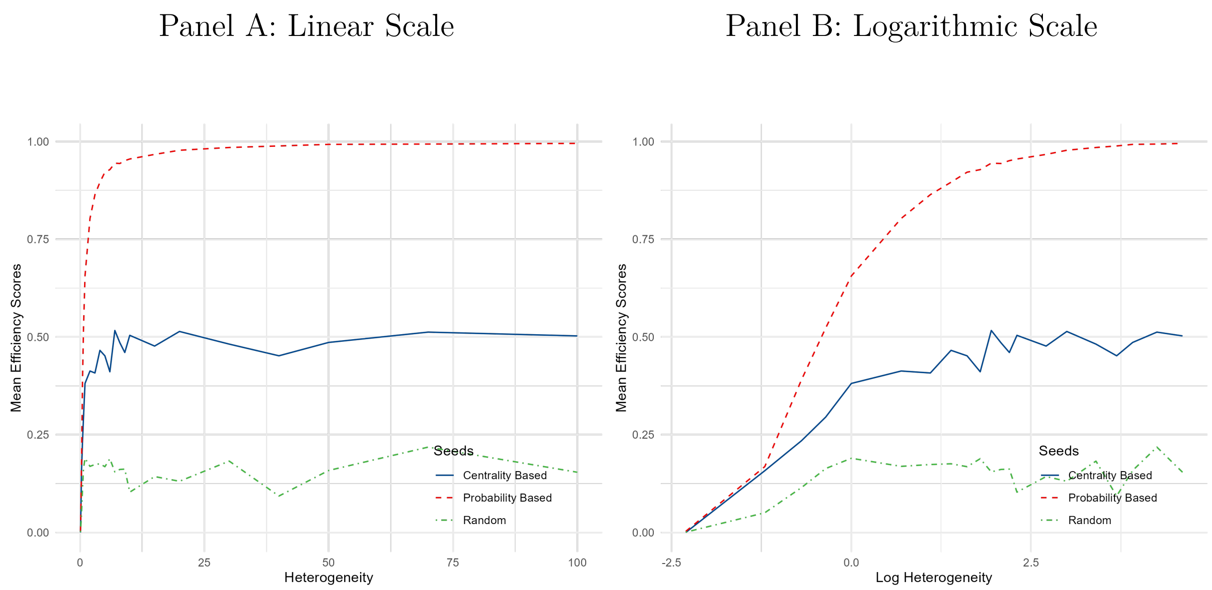


Figure G.7: Efficiency scores over increasing levels of heterogeneity (w.r.t 40 households, instead of 30, per network)

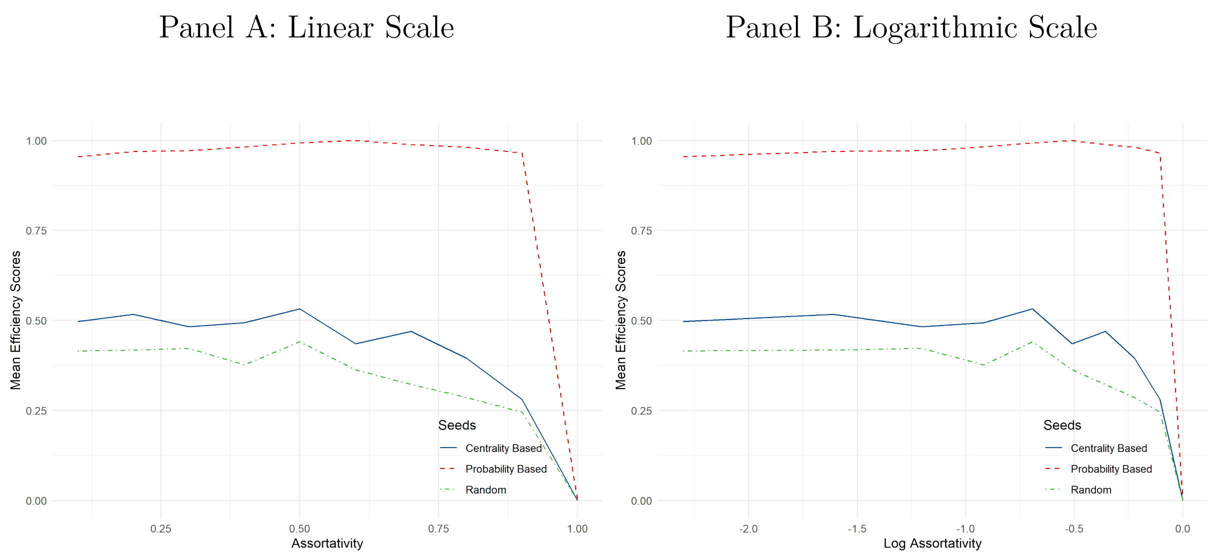


Figure G.8: Efficiency scores over increasing levels of assortativity (with same heterogeneity)

Panel A: Linear Scale | Panel B: Logarithmic Scale

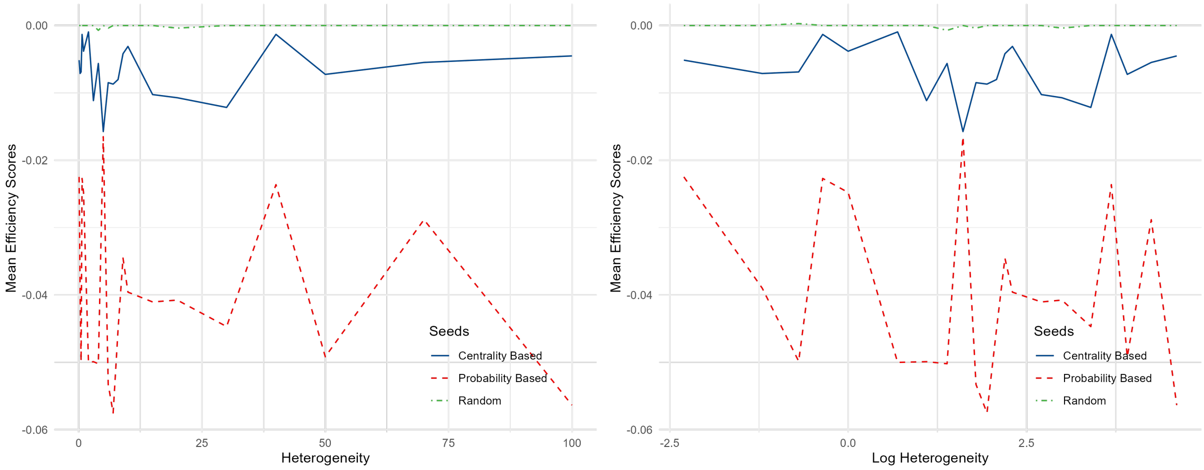


Figure G.9: Efficiency scores over increasing levels of heterogeneity (with non-assortative networks)

Table G.6: 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*** | 0.013*** | 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 G.7: 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*** | 0.102*** | 0.078*** | 0.062*** | 0.046*** |
| | (0.004) | (0.004) | (0.002) | (0.002) | (0.026) | (0.026) | (0.010) | (0.010) |
| Children | 0.006*** | 0.004** | 0.004*** | 0.002*** | 0.042*** | 0.025** | 0.020*** | 0.013*** |
| | (0.002) | (0.002) | (0.001) | (0.001) | (0.012) | (0.011) | (0.005) | (0.005) |
| Housing | 0.017*** | 0.015*** | 0.003 | 0.002 | 0.101*** | 0.102*** | 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** | 0.069*** | 0.045*** |
| | (0.004) | (0.004) | (0.002) | (0.002) | (0.024) | (0.023) | (0.011) | (0.010) |
| Assets | 0.048*** | 0.034*** | 0.022*** | 0.016*** | 0.295*** | 0.216*** | 0.125*** | 0.086*** |
| | (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 | 5608 | 5604 | 5608 | 5604 | 5608 | 5604 | 5608 | 5604 |
| pseudo R-squared | 0.248 | 0.360 | -0.168 | -0.219 | 0.027 | 0.039 | 0.032 | 0.042 |

*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 G.8: 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 | -0.092 | -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$_v$ × 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$_v$ × Heterogeneity$_v$ | 6.715** | 6.628** | 4.762* | 4.779* | 18.484*** | 19.667*** | 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 G.9: 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 | -10.419*** | -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$_v$ × Heterogeneity$_v$ | -2.423** | -2.784** | -2.237** | -2.488** | -6.692 | -10.787** | -6.574 | -9.434** |
| | (1.093) | (1.365) | (0.996) | (1.249) | (4.503) | (5.001) | (4.119) | (4.464) |
| Centrality$_v$ × Heterogeneity$_v$ × 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$_v$ × Heterogeneity$_v$ × 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$_v$ × Heterogeneity$_v$ × 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$_v$ × Heterogeneity$_v$ | 5.881** | 4.908** | 4.104* | 3.359 | 22.973*** | 16.591** | 12.347 | 8.010 |
| | (2.437) | (2.488) | (2.286) | (2.345) | (7.720) | (8.021) | (7.626) | (7.785) |
| Probability$_v$ × Heterogeneity$_v$ × 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$_v$ × Heterogeneity$_v$ × 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$_v$ × Heterogeneity$_v$ × Geo | -2.588** | -4.335** | -2.562** | -3.952** | -0.391 | -18.536*** | 0.538 | -14.549*** |
| | (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 G.10: Village level Regression 1 with Different Measure of Probability

| Variables | Adoption Rate (1) | Adoption Rate (2) | Any Non-Seed Adopters (3) | Any Non-Seed Adopters (4) |
|---|---|---|---|---|
| Eigenvector Centrality of Seeds | 0.999* | 0.817* | 0.984 | 1.067 |
| (=Centrality$_v$) | (0.565) | (0.480) | (1.303) | (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$_v$ × Heterogeneity$_v$ | -4.481* | -3.936* | -4.874 | -5.907 |
| | (2.623) | (2.281) | (6.889) | (6.437) |
| Probability$_v$ × Heterogeneity$_v$ | 10.325* | 7.276 | 23.126 | 0.889 |
| | (6.160) | (5.623) | (14.187) | (13.397) |
| 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 G.11: Village level Regression 2 with Different Measure of Probability

| | Adoption Rate | | Any Non-Seed Adopters | |
|---|---|---|---|---|
| Variables | (5) | (6) | (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$_v$ × Heterogeneity$_v$ | -4.619* | -4.617* | -12.422 | -12.190 |
| | (2.549) | (2.473) | (8.555) | (7.660) |
| Centrality$_v$ × Heterogeneity$_v$ × Complex | 1.432* | 1.595** | 9.431** | 8.099** |
| | (0.749) | (0.720) | (4.323) | (3.996) |
| Centrality$_v$ × Heterogeneity$_v$ × Simple | 0.492 | 0.576 | 3.308 | 1.958 |
| | (0.860) | (0.831) | (4.665) | (4.340) |
| Centrality$_v$ × Heterogeneity$_v$ × Geo | 3.957* | 3.711** | -1.692 | -2.661 |
| | (2.057) | (1.785) | (4.676) | (4.495) |
| Probability$_v$ × Heterogeneity$_v$ | 10.265* | 7.702 | 33.705* | 13.412 |
| | (5.561) | (5.378) | (17.388) | (16.257) |
| Probability$_v$ × Heterogeneity$_v$ × Complex | -0.316 | -0.589 | -2.606 | -1.839 |
| | (0.762) | (0.778) | (4.577) | (4.315) |
| Probability$_v$ × Heterogeneity$_v$ × Simple | 0.428 | 0.416 | 1.355 | 3.119 |
| | (0.984) | (0.866) | (5.269) | (4.868) |
| Probability$_v$ × Heterogeneity$_v$ × 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 G.12: 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$_v$ × Heterogeneity$_v$ | -1.325* | -1.020* | -1.552 | -1.997 |
| | (0.716) | (0.558) | (1.896) | (1.823) |
| Probability$_v$ × Heterogeneity$_v$ | 5.610** | 3.814 | 17.554** | 6.849 |
| | (2.660) | (2.439) | (6.873) | (6.939) |
| 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 G.13: 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$_v$ × Heterogeneity$_v$ | -1.457** | -1.181** | -2.508 | -3.114 |
| | (0.591) | (0.478) | (1.935) | (1.939) |
| Centrality$_v$ × Heterogeneity$_v$ × Complex | 0.307** | 0.304** | 1.446* | 1.355* |
| | (0.137) | (0.140) | (0.838) | (0.810) |
| Centrality$_v$ × Heterogeneity$_v$ × Simple | 0.364** | 0.395*** | -0.401 | -0.498 |
| | (0.157) | (0.152) | (0.934) | (0.917) |
| Centrality$_v$ × Heterogeneity$_v$ × Geo | 0.679** | 0.667** | 0.517 | 0.140 |
| | (0.267) | (0.262) | (0.988) | (0.914) |
| Probability$_v$ × Heterogeneity$_v$ | 4.791** | 3.306 | 19.312*** | 9.942 |
| | (2.281) | (2.166) | (7.105) | (6.963) |
| Probability$_v$ × Heterogeneity$_v$ × Complex | -0.351 | -0.419 | 0.056 | 0.189 |
| | (0.632) | (0.637) | (3.155) | (3.031) |
| Probability$_v$ × Heterogeneity$_v$ × Simple | -1.125* | -1.235* | 4.299 | 5.406 |
| | (0.664) | (0.629) | (3.876) | (3.727) |
| Probability$_v$ × Heterogeneity$_v$ × 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.