# Casting a Wider Net: Sharing Information Beyond Social Networks

Manzoor H. Dar, Alain de Janvry, Kyle Emerick, Erin M. Kelley and Elisabeth Sadoulet*

July 12, 2020

**Abstract**

Can community members be induced to share information beyond their existing social networks? Using a field experiment in Bangladesh, we show that demonstration plots in agriculture – a common technique where the first users of a new seed variety cultivate it side-by-side with an existing variety – can trigger information exchange beyond existing peer groups by signaling to others that an experiment is taking place. We compare this diffusion method with the common approach of seeding new technologies with influential farmers within the network. Using the complete social networks of about 22,000 farmers, we show that demonstration plots — when cultivated by randomly selected farmers — improve knowledge by just as much as improved seeding with influential farmers. We combine this diffusion experiment with an impact experiment to document the benefits of the new technology, and show that demonstration plots and improved seeding reach the farmers that are less likely to benefit from adoption.

*Dar: seedmanzoor.dar@gmail.com; de Janvry: University of California at Berkeley, 207 Giannini Hall, Berkeley, CA 94720-3310, alain@berkeley.edu; Emerick: Tufts University and CEPR, 8 Upper Campus Road, Medford, MA 02155-6722, kyle.emerick@tufts.edu; Kelley: World Bank, erinmkelley@worldbank.org; Sadoulet: University of California at Berkeley, 207 Giannini Hall, Berkeley, CA 94720-3310, esadoulet@berkeley.edu. We acknowledge financial support from the Standing Panel on Impact Assessment of the CGIAR and from USAID through the Feed the Future Innovation Lab for Assets and Market Access (AID-OAA-L-12-00001). The contents are the responsibility of the authors and do not necessarily reflect the views of USAID or the US Government. Emerick is grateful to the Institute of Economic Development at Boston University where he was a visiting scholar while part of this research was carried out. This paper was previously circulated under the title "Endogenous Information Sharing and the Gains from Using Network Information to Maximize Technology Adoption."

# 1 Introduction

People commonly rely on their peers for information. Research has established the existence of such peer effects across a variety of domains, ranging from learning in schools to technology adoption in low-income countries.[1] Given the importance of social networks, recent work has sought to determine who should receive new information first to ensure that it reaches as many people as possible. Proven seeding methods include mapping the full network in a community and applying diffusion models to identify optimal entry points (Beaman et al., 2018), or asking community members to identify the best individuals for spreading information (Banerjee et al., 2019). These strategies rely on the structure of the social network being approximately fixed over time, and agents passing information among their connected peers.

Another strategy — and one that has received less attention in the literature — is to ask what can be done to encourage individuals to seek information from relevant sources outside their networks with whom they may not easily connect? In this paper we investigate whether a commonly used technique in agricultural extension can effectively trigger information exchange through new interactions, and substitute for the typically difficult task of finding optimal entry points in social networks. Specifically, we focus on the use of highly visible demonstration plots that indicate the farmer is taking part in a careful scientific experiment.[2] In doing so, demonstration plots have the potential to trigger interest and induce communication beyond existing network links.

To this end, we run two related experiments spread across 256 villages in rural Bangladesh. The first experiment contrasts the two approaches detailed above to spreading information about a new rice variety called BRRI Dhan 56 (or BD56 for short). This variety can be harvested earlier than traditional rice varieties and is profitable for farmers who take advantage of its faster maturation to plant a third crop.[3] As BD56 may be suitable for some farmers and not others, we focus primarily on the spread of knowledge about BD56 rather than adop-

[1] A non-exhaustive subset of research in this area includes peer effects on academic performance (Sacerdote, 2001), purchases of financial assets (Bursztyn et al., 2014), adoption of improved sanitation in developing countries (Guiteras, Levinsohn, and Mobarak, 2015), the decision of whether to purchase crop insurance (Cai, de Janvry, and Sadoulet, 2015), and the adoption of agricultural technology (Foster and Rosenzweig, 1995; Munshi, 2004; Bandiera and Rasul, 2006; Conley and Udry, 2010).

[2] While the way demonstration plots are set-up can vary by context, they always signal that a careful scientific experiment is taking place. We focus on a particular method of demonstration whereby farmers within the community compare two rice varieties by planting them side by side on contiguous plots of land within their own farms.

[3] The variety is ready for harvest a full month before the common variety being grown in the rainy season. As a result, the technology provides enough time for a third crop to be grown in between the rainy and dry-season rice crops. This large change in the cropping system — going from producing two rice crops to producing a third crop in between — enhances the potential for learning about BD56 and its attributes.

tion rates. We introduced the new variety to five farmers, referred to as “entry points”, in a random subset of 192 villages. Villages were cross randomized across two treatments: (1) the selection criteria for these entry-point farmers and (2) the demonstration method, aimed at encouraging other farmers from the village to seek information from the demonstrating farmer. Figure 1 shows the different treatment cells.

In terms of selection of entry points, we randomized villages across three different methods. In 64 villages farmers were randomly selected (as a benchmark). In the next 64 villages we relied on the local knowledge of agricultural extension officers (known as “sub-agricultural officers” or SAO’s) to identify farmers who would be most effective at demonstrating a new rice variety. In the remaining 64 villages, we ranked farmers according to farm size and selected the five largest farmers.[4] The improved selection of entry points – using large or SAO designated farmers rather than random – seeks to test practically feasible approaches to finding the most influential farmers in the network. Importantly, the notion of optimal entry points is designed to exploit the network as it exists at baseline, and does not consider that the intervention itself may change the network.

Turning to the demonstration method, we randomized whether the entry point farmers were asked to set up demonstration plots or not. In 32 out of the 64 villages assigned to a particular entry-point selection arm, farmers were assisted in setting up “head-to-head demonstration plots”, which involved cultivating BD56 alongside a counterfactual seed variety of the farmer’s choosing. We provided two markers — one reading “BD56” and the other listing the name of the chosen alternative variety — to make comparison across the two varieties highly visible (see Figure 2 for an example). This method of demonstration signals that a scientific experiment is taking place as it resembles methods employed by agricultural extension agents when they want to demonstrate the attributes of a new seed variety.[5] In the remaining 32 villages, the entry points were provided with a single “BD56” marker for their BD56 plot. Demonstration plots do more than broadcast information about the existence of this new technology, because entry points in all arms labelled their fields with a BD56 marker. Instead, they show the existence of a careful experiment that can be discussed with the demonstrating farmers.

[4] We found during piloting that large landholders often act as opinion leaders during focus group discussions. Moreover, other farmers seemed to look at large farmers to learn about new technology. These casual observations led to the inclusion of the arm where farm size was used to select entry points.

[5] Demonstration plots set up by agricultural extension agents typically take place in large clusters of contiguous plots each with a particular technique to allow for comparisons (“Cluster demonstrations”). In this study, we bring a variation of this method to the village level by inviting a large number of individual farmers to set up small comparative demonstration plots on their own land (sometimes referred to as “comparative demonstration plots” or “head-to-head demonstration plots”. We didn’t provide any further assistance, such as advice or inputs, with the actual cultivation of the two plots. This was purposeful to let the farmers adapt the technology to their own circumstances and make the approach easily scalable.

This first experiment delivers three main results. First, we verify that the large and SAO-selected entry points are far more central in the village networks than random farmers, and are thus well positioned to spread information. At baseline, we conducted a full census of the 256 villages in our sample (22,000 surveys).[6] Using these data, we observe that the average entry point in the random villages is connected to 4.6 other farmers. This increases sharply to 8.2 and 9.1 connections for entry points in SAO and large farmer villages, respectively. Similarly, the eigenvector centrality of entry points increases by 47 percent under SAO selection and 80 percent with large-farmer selection.

Building on this result, we show that seeding with large and SAO-selected farmers improves knowledge transmission. We focus on knowledge transmission as our primary outcome because it was only profitable for some farmers to adopt BD56.[7] We measure knowledge by conducting a survey with 10 random farmers in each village after the harvest of BD56. We ask farmers whether they know about BD56 and its basic attributes, and the number of conversations they have had about BD56. Using large farmers as entry points increases the rate of being informed by 7.4 percentage points (12.3 percent) in villages without demonstrations. Similarly, entry points selected by extension agents increase knowledge by 6.7 percentage points (11.2 percent) in villages without demonstrations.

Second, we find that these gains from improved seeding can also be obtained using demonstration plots with *random farmers*. Demonstration plots increase the rate of being informed by 7.2 percentage points (12 percent) with random entry points. Noticeably, the magnitude of this effect is similar to what we find with the improved-selection of entry points. The results remain similar when looking at the number of reported conversations about BD56: demonstration plots with random farmers induce conversations by about the same amount as introducing seeds with more central farmers. This finding offers insight into how interventions that broadly attract attention can substitute for seeding with more central entry points in networks.

Third, we proceed to investigate what explains the effectiveness of our treatments. Starting with the demonstration plots, we focus on the 64 villages where entry-points were chosen randomly, and peer effects can be causally identified. We find that an additional connection with an entry point has no effect in demonstration villages, but increases knowledge by 13.5 percentage points (22 percent) in non-demonstration villages. This suggests that demon-

[6] Chandrasekhar and Lewis (2016) show that measures of network centrality are misleading when estimated using only a sample of nodes within the network. Our approach of fully characterizing the network by surveying each household in the village alleviates this concern.

[7] BD56 was only profitable for farmers who significantly changed their cropping cycles by planting a third crop. Note we also observe uptake when BD56 was made available for sale at subsidized prices. The results on seed adoption are noisier, but qualitatively consistent with our observations on knowledge detailed here.

stration plots entirely eliminate peer effects. In other words, demonstration plots lead to information exchange outside of baseline networks and therefore induce broad transmission of knowledge. Also consistent with this network interpretation, we find that demonstration plots were most effective for farmers that are least connected in the network — where connectivity is measured by eigenvector centrality. A plausible explanation is that demonstration plots induced these less connected farmers to endogenously seek information.

In contrast, we show that information diffusion via existing links partly explains the effectiveness of seeding with large and SAO-selected farmers. We find that the effects of large and SAO selection on knowledge decrease by 43 and 31 percent, respectively, when conditioning on the average degree centrality of entry points. Conditioning on average degree centrality is not a perfect test — since the particular network measure that should "knock out" the treatment effects depends on the specific diffusion model.[8] Nonetheless, the result is consistent with the idea that diffusion via network links partly explains why the entry-point treatments were successful.

We turn to our second experiment to identify the benefits of the new technology, and test whether demonstration plots (or alternative seeding technologies) have a higher probability of delivering information to those farmers who are most likely to reap the benefits if they adopt. Returning to Figure 1, we randomly selected 64 control villages where we provided a long-duration rice seed (known as BRRI Dhan 51 - BD51 for short) to a set of farmers identified using the same criteria as in the 192 BD56 villages. Using these farmers as controls, we find that the main benefit of the short-duration BD56 seed is the ability to increase cropping intensity by growing a third crop in between the rainy and dry-season rice crops. We use the machine learning methods developed in Chernozhukov et al. (2018) to estimate a mapping between baseline covariates and the treatment effect of BD56 on the number of crops grown. This "predicted benefit index" serves as a prediction of which farmers are most likely to benefit from BD56 by increasing cropping intensity.

Using this predicted benefit index in the 192 villages of our first diffusion experiment, we find suggestive evidence that both demonstration plots and our selection treatments only increase knowledge and conversations for farmers that have *below-median* expected treatment effects of BD56 on the number of crops grown. Put differently, those most likely to benefit from a new technology find a way to become informed even when entry points are chosen arbitrarily and demonstration plots are not used. These findings point to an important consideration for research that studies alternative mechanisms for increasing diffusion of

[8] Degree is a suitable measure if we think of diffusion models with few periods, the probability of passing information to connected friends is high, and the farmers with the largest degrees are sufficiently spread out in the network.

products that have heterogeneous benefits. Combining diffusion experiments with standard impact evaluations allows the researcher to estimate treatment-effect heterogeneity and use that to measure whether different diffusion strategies reach the people who are most or least likely to benefit from an innovation, even without a strong prior on which observables drive the heterogeneity.[9]

The last part of the paper shows how a simple diffusion model, when amended to allow for formation of new links with entry points, explains the pattern of our results. In the model, farmers can either become informed by receiving information flowing through the structure of existing links, or by actively communicating with entry points. Farmers in the worst position to learn from the network, i.e. those that are least connected to entry points, benefit most from having the opportunity to communicate directly with entry points. The demonstration plot treatment appears to deliver these benefits by signaling to others that an experiment is taking place, which induces unconnected and more isolated farmers to seek information from entry points directly.

Our experiment contributes to the literature on endogenous communication among community members. Theoretical models have been developed in which agents need to invest in creating costly communication links in order to acquire information (Acemoglu, Bimpikis, and Ozdaglar, 2014; Calvó-Armengol, Martí, and Prat, 2015). Empirical work in this area has focused on two areas. First, how do people aggregate information across sources? Both Mobius, Phan, and Szeidl (2015) and Chandrasekhar, Larreguy, and Xandri (2020) use lab experiments to distinguish between different models of information aggregation in networks.[10] Second, how should information be introduced into a community to maximize its spread? The answer may depend on the types of frictions affecting learning. Two papers explore one particular friction: seeking information can reveal low skill and create a stigma effect (Chandrasekhar, Golub, and Yang, 2019; Banerjee et al., 2018a). In this setting, it may be optimal to seed information with a few individuals and make it broadly known that these people are informed. The literature has yet to consider different ways to induce communication between agents after entry-points have been selected. We do so by fostering the emergence of new communication links through demonstration plots and showing that farmers can be induced to seek information and engage in conversations beyond their pre-existing network.[11]

[9]Rigol, Hussam, and Roth (2017) is the closest example where the authors use machine learning methods to estimate treatment effect heterogeneity for microfinance in India. They then show that using community information on the returns to microfinance is more effective than the machine learning algorithm when applied to the set of observables in their baseline data.

[10]These studies use games whereby the experimental subjects are given noisy information about the state of the world, and can win a prize for finding out the true state. In such a setting where all agents have equally noisy signals, the choice of whom to talk to is solely driven by costs.

[11]Empirical work by Banerjee et al. (2018b) emphasizes the importance of endogenous network formation,

Turning to magnitudes, our experiment compares this method of inducing communication with the more popular approach of seeding information with influential people. These seeding studies address the demonstration of agricultural inputs (Beaman et al., 2018; Beaman and Dillon, 2018), the diffusion of information about microfinance (Banerjee et al., 2013), diffusion of health products (Kim et al., 2015), and information on how to capitalize on a financial opportunity or the uptake of vaccines (Banerjee et al., 2019). There are some limitations with relying on this approach: it can be costly, and the efficacy of finding better entry points for seeding information likely depends on the underlying structure of the social network or the specific model of diffusion (Centola, 2010; Valente, 2012; Golub and Jackson, 2012).[12] Rather strikingly, our findings suggest that signaling the availability of new information through demonstration plots results in diffusion rates similar to those achieved with the best entry points.[13]

Focusing on agricultural development and policy, information frictions have been shown to be one of the reasons why farmers do not adopt new technologies.[14] Agricultural extension is expected to serve as the policy tool to improve learning. In practice, agricultural extension relies mostly on seeding information with selected farmers to diffuse information at the village level. It also uses large scale cluster demonstration sites where extension agents have groups of farmers demonstrate alternative techniques or varieties under their direction with the provision of free inputs. Our experiment borrows from these concepts, showing that the establishment of farmer-level demonstrations indicates that careful experimentation is taking place within the village, thereby inducing communication within the community.

The remainder of the paper is organized as follows. Section 2 discusses the design and implementation of the experiment, data collection, and basic characteristics of the sample. Section 3 presents each of our individual results, focusing on how demonstration plots influence learning, and on understanding what drives their effectiveness. Section 4 shows how the impact of BD56 varies across farmers and tests whether the treatments deliver information to farmers with higher predicted benefits. Section 5 outlines a simple theoretical framework

but focuses on a more drastic restructuring of the network rather than on the one-time connections we focus on here.

[12] Banerjee et al. (2019) show how to overcome this difficulty by asking a sample of villagers who are the important people for diffusing information. Akbarpour, Malladi, and Saberi (2018) shows that in many network structures the benefits from seeding information with a slightly larger number of agents outweigh the benefits of identifying the most central individuals.

[13] Other work has looked at different ways of improving the contact farmer model, including compensating entry points for effort (BenYishay and Mobarak, 2019) and training them (Kondylis, Mueller, and Zhu, 2017). The H2H method of attracting attention represents a cost-effective alternative, only requiring that two sticks be planted in farmers' fields to signal information.

[14] In addition to information, numerous studies highlight a wide range of explanations, including behavioral biases, profitability, and risk (Duflo, Kremer, and Robinson, 2011; Suri, 2011; Karlan et al., 2014; Emerick et al., 2016; Cole, Giné, and Vickery, 2017).

that is consistent with our results. We then provide an overview and discuss implications of the findings in the final section.

# 2 Overview of the Experiment

We conducted the study in 11 sub-districts (upazilas) scattered across 3 districts of Rajshahi division, consulting with the Department of Agricultural Extension to identify upazilas that were suitable for the rice variety being introduced.[15] We sampled 23 or 24 villages per upazila, focusing on those with no more than 150 households, resulting in a final sample of 256 villages. This includes the 192 villages that received the new BD56 rice variety in the diffusion experiment, as well as 64 control villages that received the BD51 longer duration rice variety for the impact evaluation experiment.[16] This village-level randomization was stratified by upazila.

## 2.1 Experimental design

Figure 1 summarizes the experimental design for the diffusion and impact evaluation experiments. For the diffusion experiment, 192 villages were first subdivided into three groups of 64 villages each. In the first group of 64 villages, the seeds were distributed to five farmers selected at random. In the second group, we ranked farmers by landholding sizes and provided BD56 seeds to the top five farmers. In the third group of 64 villages, we asked the Sub-Agricultural Officer (SAO) to identify five farmers in the village that would be effective at demonstrating the new variety. Within each group of 64 villages, we then selected 32 villages to receive additional assistance setting up H2H demonstration plots. We provided two sticks to farmers: one with the name of the new variety (BD56), and another with the name of the variety they had selected to plant alongside. Farmers in the remaining 32 "non-demonstration" villages received a single sign for their BD56 plot. The provision of the single sign ensures that any effect we detect in the demonstration plot villages with two signs goes beyond the attention effect of placing one sign in the field.

We selected H2H demonstration plots as they signal to villagers that a trial is taking place, thereby generating interest and inducing interactions. Farmers are familiar with head-to-

[15] See Figure A1 for the location of the 11 upazilas included in the study.

[16] This variety was chosen due to its similarity to the most popular variety at baseline. BD51 is released as "Swarna-Sub1" in India and several other countries. Emerick et al. (2016) show that Swarna-Sub1 is similar to Swarna, besides Swarna-Sub1 being more flood tolerant. However, our sample is not a flood prone area. We introduced Swarna-Sub1 in the control villages because Swarna is not officially released in Bangladesh (and thus not available for sale) despite the fact that 77 percent of farmers in the village census reported growing Swarna at baseline.

head experimentation as agricultural extension agents often set up large scale demonstrations plots, and organize information sessions for farmers in the surrounding area to attend. We adapt this model by selecting different types of farmers to showcase the new variety on smaller plots of land within their own village, and let the interactions between farmers happen organically.

For the impact evaluation experiment, we selected up to 15 "counterfactual" farmers to receive equal sized amounts of BD51 seed in each of the 64 control villages. This included five farmers with the largest landholdings, five farmers selected by the SAO, and five farmers selected at random. These sets overlapped in some cases and therefore the number of farmers per control village is occasionally less than 15. Identifying these particular farmers in control villages was necessary in order to compare how the entry points we identified in the treatment groups cultivated BD56 relative to the longer duration counterfactual we distributed in the control villages.

## 2.2 Timeline and data collection

*Census with network information of each village*

Figure 3 presents the timeline for the study. We began by performing a complete census of each village in March 2016. Villages had on average 86 households, resulting in a population of 21,926 households. We collected information about farmers' annual agricultural production (landholdings size, fertilizer use, production, and varieties sown). There are up to 3 cultivation periods in Bangladesh: Aman, Rabi and Boro — where Aman and Boro are dominated by rice. We also captured information about their social networks, the name of the person they considered to be the best farmer, and the names of up to 10 farmers they talked to about rice farming during the last Aman season. We use these data to identify the largest farmers in each village, select the random entry-points, compute network statistics, and predict heterogeneous impacts of BD56 as a function of observable covariates.

Table A1 presents summary statistics from the household census and verifies randomization balance.[17] Our sample consists primarily of farmers cultivating long-duration rice varieties in the Aman and Boro rice seasons (only 1.17% of treatment farmers and 2.4% of control farmers planted short-duration varieties in the 2015 crop cycle). While approximately 35% of farmers only grow rice in two seasons, a non-negligible share grow a Rabi crop (including wheat, potato, pulses, onion, and garlic) between the two rice seasons. Finally, farms in the sample are small. Average area sown with Aman rice (the main crop) is approximately 1.33 acres.

[17] Table A2 shows randomization balance for the sample of entry points in the impact experiment.

We define two farmers as being connected or being peers if either one names the other. We use this information to create various centrality statistics including degree centrality (the number of connections a person has), eigenvector centrality, and betweenness centrality (the number of times a node acts as a bridge along the shortest path between two other nodes). Figure A2 displays the distributions of these three measures for the entire set of households interviewed during the census. Farmers on average talk about rice with 4 of their peers, though some interact with up to 26. The distributions of eigenvector and betweenness centrality display similar patterns: most farmers have a few connections, while some have a disproportionately high number. These long right tails suggest that farmers' information networks feature some highly central farmers that could serve as more effective entry points.

*Seed distribution*

In June 2016, we distributed 5kg minikits to 1,795 entry-point farmers.[18] During these visits, the field team carefully explained the features of the seeds being distributed, and provided farmers with calendars to record the dates crops were sown, harvested, irrigated and applied with inputs. We also supplied farmers with sticks and cards to place in their fields as a way of demonstrating to other farmers the variety they were planting. We visited farmers 6 weeks later to make sure we answered any remaining questions they had about the seeds, and to verify that the sticks were properly displayed in the fields.

*Information diffusion survey*

In April 2017, we visited all treatment villages again. We surveyed 10 additional farmers, selected at random, to determine if they had heard about the BD56 seeds that were introduced to entry points 10 months earlier. The survey also asked whether they knew about the variety's key features. We use this information to assess whether farmers' knowledge about BD56 differs by treatment status. We focus on knowledge transmission as our primary outcome because the technology was not necessarily profitable for all farmers to adopt. We will see that positive returns could only be reaped by farmers who took advantage of BD56's short-duration to plant an additional crop.

*Seed sale*

In June 2017, we visited each treatment village to sell 5kg and 2kg bags of BD56 seeds at a 60% discount. The field team called a select sample of farmers in each village to inform

[18] The total number of entry points is 5 per BD56 village, 100% of whom were reached. The total number of entry points is up to 15 per BD51 village (though there were occasionally less due to overlapping sets), 99% of whom were reached.

them about the date and time of the seed sale. The sample included the original minikit recipients, and the ten randomly selected farmers who had been surveyed about their BD56 knowledge 2 months prior. The field team travelled to each village on a pre-determined date, and set-up their truck in the middle of the village (often at the local market) with a large sign and recorded each sale in a tablet. While we did not record the identity of the buyer, the survey provides a measure of BD56's diffusion within the village. Unfortunately, we ran out of seeds before reaching all of the villages, and hence this information is only available for 168 of the 192 villages.

*Agricultural surveys of entry-points.*
In addition to the main sources of data described above, we conducted a series of agricultural surveys with the entry-points in all 256 villages. The goal of these surveys was to rigorously establish the impact of BD56 on agronomic practices, cropping intensity, and annual income. A baseline survey was administered at the time of the minikit distribution in June 2016. We asked farmers to provide this information for 3 of their plots, which we selected randomly when farmers had more than 3. We conducted three additional survey rounds in order to fully characterize annual production. This includes visits in January 2017 to capture production of the Aman rice crop (midline 1); in April 2017 to document Rabi crop production (midline 2); and in August 2017, to record crop production levels during the Boro season, and crop choice for the 2017 Aman rice season.

# 3 Results: Diffusion Experiment

We now present our main results from the diffusion experiment. We start by characterizing differences between entry points in their network centrality. We then compare the gains in information diffusion from using these more central entry points with those from H2H demonstration plots. Additional results suggest that H2H demonstration plots are effective because they induce more communication. In contrast, a standard network diffusion model explains much of the effectiveness of our selection treatments.

## 3.1 The network centrality of entry points

Compared to eliciting the entire network, our two methods of selecting entry points are less demanding in terms of data: entering with large farmers requires only administrative data on farm sizes and SAO-based selection requires only a short interview with an agricultural extension agent. We first check that these two methods deliver entry points that are theoretically good for diffusion using our social network survey.

Measures of network centrality differ noticeably between random entry points, those identified by SAO's, and the largest farmers. Table 1 displays average characteristics for entry point farmers in BD56 treatment villages. The average random farmer has about 4.6 connections with other farmers in the village. The entry points selected by SAO's have an average of 8.14 connections, and the largest farmers have an average degree of 9.04. Eigenvector centrality of entry points increases by 47 percent with SAO selection and 80 percent for the largest farmers. Banerjee et al. (2013) introduce the concept of diffusion centrality of farmer $i$ as a measure of the expected number of times that farmers obtain a piece of information that was introduced with $i$.[19] Compared to random entry points, average diffusion centrality of SAO-selected entry points is higher by 0.64 standard deviations and the diffusion centrality of large entry points increases by 0.87 standard deviations.[20] Figure A3 shows the cumulative distribution functions of the different network centrality measures across the selection treatments. Most importantly, the increases in centrality for large and SAO farmers occur throughout the distributions. Moreover, seeding with large farmers generates entry points that are around the 90th percentiles of the degree and eigenvector centrality distributions (the 90th percentiles of the degree and eigenvector centrality distributions are 9 and 0.157, respectively). Moving from the median to the 90th percentile of the centrality distribution should generate meaningful improvements under standard models of peer-to-peer learning. These findings demonstrate the relevance of our selection treatments as benchmarks.

Table 1 also shows how several other observable characteristics vary across entry point treatments. There are two notable observations. First, SAO-selected entry points tend to be larger farmers, with farm sizes that are about 65 percent larger than those of random farmers. We show in Table A3 that controlling for farm size reduces substantially the gaps in network centralities between SAO-selected and random farmers. Put differently, extension officers could be using knowledge of farm size in selecting entry points, and this explains part of the reason why SAO-selected entry points are more influential in networks.[21] This phenomenon

[19]This measure requires as parameters the number of periods for the diffusion process and the probability that an informed agent passes information to their social connections. We set the number of periods to 5 and the information passing probability to 0.5. We also normalize the measure by subtracting the village-specific mean and dividing by the village-specific standard deviation.

[20]The high centrality of large and SAO farmers — relative to random entry points — is comparable to other mechanisms to selecting entry points that have been tested in the literature. For instance, Banerjee et al. (2019) find that the median diffusion centrality of people reported as suitable for spreading information is larger than that of other villagers by around 0.5 to 1 standard deviations. We find that relative to random targeting, larger farmer targeting would increase median diffusion centrality by around 0.51 standard deviations (Figure A3).

[21]The ability of extension agents to select influential entry points contrasts with Beaman et al. (2018) who find that Malawian extension agents possess little information on the optimal entry points within social networks. The sharp correlation between farm size — an easily observable characteristic — and network

is visually evident in Figure 4 where we show the network structures for 6 randomly selected villages with either SAO or large-farmer selection. All three of the large-farmer villages in the top panel have at least one relatively larger farmer that is well connected in the network. Focusing on the SAO-selection villages in the bottom panel, both village 99 and 224 have medium or larger size farmers that are central in the network and were selected as entry points by the SAO.

Second, our door-to-door census asked each household to list the "best farmer" in the village. A random farmer is only named about 0.79 times while SAO entry points are named 5.27 times, and the five largest farmers in each village 6.4 times. These numbers further suggest that our selection treatments identify entry points that are both better networked and that other farmers consider to be knowledgeable about agriculture.[22]

## 3.2 Diffusion of information across treatments

Do H2H demonstration plots make people more informed about the existence of a new technology? If so, how do these effects compare to those generated by improved selection of entry points? We compare knowledge across the six different arms of the diffusion experiment. The corresponding regression is:

$$\begin{aligned} informed_{ivs} = \beta_0 + \beta_1 RandomDemo_{vs} + \beta_2 SAONoDemo_{vs} + \beta_3 SAODemo_{vs} \\ + \beta_4 LargeNoDemo_{vs} + \beta_5 LargeDemo_{vs} + \alpha_s + \varepsilon_{ivs}, \end{aligned} \tag{1}$$

where the dependent variable is an indicator for whether farmer $i$ in village $v$ and upazila $s$ is informed about BD56, $SAONoDemo_{vs}$ is an indicator for villages with SAO selection and no H2H demonstration plots, and the remainder of the variables are defined analogously. The dependent variable is an indicator for whether the farmer has heard of BD56. We found that farmers who know about BD56 also knew the basic properties that it has a shorter growing season.[23] Hence, our measure of being informed also includes knowing these basic properties.[24] We always include strata (upazila) fixed effects and cluster standard errors at the village level.

The results in Table 2 deliver three insights. First, H2H demonstration plots increase

centrality offers one possible explanation for the greater ability of extension agents in our sample.

[22]The number of nominations as the best farmer is correlated with network centrality. It explains 44 percent of the variation of degree centrality, 32 percent for eigenvector centrality, and 41 percent for betweenness centrality.

[23]87.39% of farmers who heard about BD56 selected "maturation period" as the variety's main benefit among a list of 7 potential benefits.

[24]We refer to the outcome as "being informed" or "having knowledge" interchangeably throughout the analysis.

knowledge when cultivated by randomly selected farmers. Specifically, the rate of being informed increases by 7.2 percentage points (12 percent) when random farmers grow BD56 side by side with a chosen comparison variety. The large rate of being informed in *RandomNoDemo* villages (60 percent) shows that information diffuses, even under the benchmark where entry points are random and demonstration plots are absent.

Second, the effect of H2H demonstration plots with random farmers is roughly of the same magnitude as the effects of entering with large and SAO farmers. As we would expect based on their network connections, entering with large and SAO-selected farmers increases the spread of information. Amongst non-demo villages, SAO selection increases being informed by 6.7 percentage points (11.2 percent) and entering with the largest farmers increases it by 7.4 percentage points (12.3 percent).[25] These effects are quite similar and statistically indistinguishable from the effect of H2H demonstration plots with random entry points. Moreover, H2H demonstration plots eliminate the effects of targeting more central farmers. Specifically, the estimates of $\beta_1$, $\beta_3$, and $\beta_5$ are statistically indistinguishable.

Third, the H2H demonstration plots have no effects when cultivated by the better-connected large and SAO farmers. The estimates of $\beta_2$ and $\beta_3$ are nearly identical, meaning that the demonstration plots didn't spread knowledge with SAO selection. Similarly, the estimates of $\beta_4$ and $\beta_5$ indicate that adding H2H demonstration plots had no effect with large-farmer entry points. The results in the next section offer a plausible explanation. Namely, H2H demonstration plots and improved seeding are substitutes because they both cause information to spread to the same types of people. The way of reaching these people, however, differs between the two techniques. Improved seeding causes information to flow through to people that were the least likely to learn from seeding with arbitrary farmers. These are the farmers that are the least central in the network. Demonstration plots draw attention and signal an experiment is taking place, thereby inducing communication with lesser known farmers.

Columns 2 and 3 of Table 2 show effects on the number of conversations farmers had about BD56.[26] While somewhat noisier, these data are also consistent with the demonstra-

[25] The effectiveness of large-farmer selection should be considered as an intention-to-treat effect because large farmers were sometimes selected as entry points in random and SAO villages. At least one of the five largest farmers was selected as an entry point in 19 of the 64 random villages and 38 of the 64 SAO selection villages. Table A4 shows that consistent with the above findings, knowledge is higher in villages where there was at least one large entry point. In addition, this effect goes away in villages with demonstration plots.

[26] Focusing on the mean outcomes, the average respondent in the random no-demo villages reported 0.84 conversations about BD56, 0.72 of which were with the entry points, and by difference 0.12 were with any of ten other randomly selected farmers that each farmer was asked about. Importantly, the reported conversations should not be interpreted as the total number of conversations about BD56, but rather the number of conversations with the 15 farmers asked about in our survey - five entry points as well as 10 randomly selected other farmers for each village.

tion plots creating just as many conversations as the improved selection of entry points. The demonstration plots led to 0.12 more conversations per farmer when entry points were selected randomly (col. 2), almost all of them being with entry points (col. 3). This 14 percent effect is similar to the effect on knowledge reported in column 1. The data on conversations help to rule out an alternative explanation whereby the side-by-side comparison — and two markers in the field — was more effective at simply broadcasting information on the existence of BD56. Instead, the demonstration plots caused farmers to engage in some form of additional information exchange, rather than just learning about the new technology from the sign in the field.

### 3.3 How do H2H demonstration plots work?

We highlight one mechanism which makes the H2H demonstration plots work: they encourage people to approach farmers outside of their immediate network to understand more about the new technology. The 64 villages with random entry points allow us to consider this mechanism. Within these villages, the number of entry points in a farmer's network is as good as randomly assigned when conditioning on their total number of connections.[27] As a result, the average effect of being connected to an additional entry point can be estimated with:

$$informed_{ivs} = \beta_0 + \beta_1 Entry\ Point\ Peers_{ivs} + \beta_2 Total\ Peers_{ivs} + \beta_3 Entry\ Point\ Peers_{ivs} * Demo_{vs} + \beta_4 Total\ Peers_{ivs} * Demo_{vs} + \beta_5 Demo_{vs} + \alpha_s + \varepsilon_{ivs},$$

where $Entry\ Point\ Peers_{ivs}$ is the variable measuring how many of the five entry points farmer $i$ is connected to and $Total\ Peers_{ivs}$ is the network degree of farmer $i$. Our hypothesis is that demonstration plots should make social relationships less important for being informed.

Ignoring heterogeneity across H2H demonstration plots, the average peer effects are indeed consistent with information exchange. Column 1 in Table 3 shows that being connected to an additional entry point increases knowledge by 7.7 percentage points, i.e. about 12 percent. More interestingly, the second column shows that connectedness with entry points makes people more informed *only* in villages without demonstration plots. An additional connection to an entry point increases knowledge by 13.5 percentage points without demonstration plots, but this effect goes down significantly to only 1.9 percentage points with demonstration plots. In addition, H2H demonstration plots only increase learning for farm-

[27]Miguel and Kremer (2004) use a similar strategy when estimating spillover effects from deworming in Kenya.

ers that were unconnected to entry points at baseline. H2H demonstration plots increase knowledge by 11.3 percentage points for farmers having no baseline connections to entry points. The effect disappears with just one connection to an entry point as the coefficient on the interaction term (11.6) is nearly identical to the coefficient on the demonstration villages indicator (11.3). Finally, the similarity of the effects points to how H2H demonstration plots substitute for social connections to entry points: the effect of H2H demonstration plots (11.3) is nearly the same as the peer effects in non-demo villages (13.5). Measuring connections with a binary variable for being connected to at least one entry point does not change the results (columns 3 and 4).

H2H demonstration plots also induced conversations between entry points and farmers that were outside of their baseline information networks. Table 4 shows analogous results where the dependent variable is instead the number of reported conversations between respondents and entry points. Results are consistent with those on knowledge. Namely, H2H demonstration plots created an environment that encouraged farmers to talk to people they otherwise would not have conversed with.

As further evidence that H2H demonstration plots induce learning, we show that they were only effective for the least networked farmers. We investigate this by limiting the analysis to the random-entry-point villages and estimating:

$$informed_{ivs} = \beta_0 + \beta_1 Demo_{vs} + \beta_2 Eigenvector_{ivs} + \beta_3 Eigenvector_{ivs} * Demo_{vs} + \alpha_s + \varepsilon_{ivs}, \quad (2)$$

where $Eigenvector_{ivs}$ is the baseline eigenvector centrality of farmer $i$. The estimated coefficient on the interaction term $\beta_3$ measures whether H2H demonstration plots had a differential effect for more central farmers. Intuitively, the more connected farmers have a number of ways (both direct and indirect) to find out about BD56. In contrast, the least central farmers have more to gain from making new connections to entry points.

Table 5 shows that the effect of H2H demonstrations varies significantly according to the farmer's eigenvector centrality. The coefficient on the interaction term ($\beta_3$) is negative and precisely estimated. The effect of H2H demonstrations is close to zero and statistically insignificant for the most central farmers, but is large and significant for the least central farmers.[28] Put another way, H2H demonstration plots are most effective for the least central farmers in the network. These farmers were in the worst position to learn with random seeding.

This evidence offers an explanation for why the identities of entry points become irrelevant

[28] The effect at the 90th percentile of the distribution is -0.022 (p=0.71) and the effect at the 10th percentile of the distribution is 0.165 (p=0.003). Figure A4 shows that the effect of the demonstration plots comes from the bottom three quartiles of the eigenvector centrality distribution.

when they cultivate H2H demonstration plots. Demonstration plots effectively encourage farmers to pay attention to others, and eliminate the need to rely on central nodes in the network to spread information.

### 3.4 How do influential farmers transmit information?

Any network-based model of diffusion with an exogenous network would favor targeting more central entry points. As a result, the effects of SAO and large-farmer selection would be predicted to decrease when conditioning on the average centrality of entry points. The data show exactly this. Table 6 shows that conditioning on the average degree centrality of entry points (moving from column 1 to 2) causes the effects of large farmer and SAO selection to decrease by 43 and 31 percent, respectively.[29] In addition, the average degree of entry points is strongly correlated with farmers having information. This relationship is consistent with, although does not prove, a network based model. There could be unobserved correlates of degree centrality that actually explain the effectiveness of the large farmer and SAO selection. The exercise instead offers evidence that the ability of large and SAO entry points to increase knowledge is associated with their more central network positions.

The main benefit of BD56 is the shorter duration and hence the ability to grow three crops during the year instead of two. As we show below, large and SAO-selected farmers were more likely to respond to the treatment by growing this additional "rabi" crop after their first rice harvest. Put differently, large and SAO-selected farmers were more likely to showcase the main benefit of BD56. This appears to have captured the attention of others and itself explains part of the reason why people are more informed with large and SAO-selected farmers. Column 3 in Table 6 shows that the entry-point treatment effects also decrease when conditioning on the number of entry points in a village that grew rabi crops on their BD56 plots. Conditioning on both this and degree centrality causes the effects of large farmer and SAO selection to decrease by 63 and 51 percent, respectively (col 4). While this does not imply causality, it suggests that SAO and large farmer's centrality, and their ability to showcase the benefits of the new variety, are highly correlated with information exchange.

In line with the network-based diffusion mechanism, Table A5 shows that seeding with large and SAO-selected farmers works for the *least central* farmers in the social network. This is especially true for eigenvector centrality — much like we found with the effectiveness of H2H demonstration plots. Conversely, seeding with random farmers works for the more central farmers: we see a positive correlation between eigenvector centrality and having

[29]We limit the data to the non-demo villages for this analysis since the selection treatments are only effective in these villages.

information when entry points are randomly selected and there are no demonstration plots.[30] Learning from peers can explain this result. Being in a better position to learn from random farmers leads to smaller gains from improved seeding strategies. In contrast, the least central farmers gain more from improved seeding because they learn the least from random entry points.

These heterogeneity results across both approaches show that H2H demonstration plots and improved seeding work for similar types of people: the least central farmers in the network. However, our evidence suggests that the mechanism for reaching these least central farmers is potentially different. Demonstration plots induce conversations and break down the role of peer effects. Improved seeding, on the other hand, facilitates peer-to-peer learning for two reasons. First, it diffuses information to farmers that are the least likely to learn from arbitrary selection of entry points. Second, more central farmers are themselves more likely to demonstrate technology in a way that captures attention.

## 3.5 Effects on seed purchases

While being informed about BD56 is our main outcome variable, we also collected data on purchases of the seeds.[31] Table 7 shows regression results akin to Equation (1), but at the village level where the dependent variable is either the number of farmers purchasing or the adoption rate (number of buyers divided by village size). The point estimates are much noisier, but the coefficients are sizable and the directions line up with what we observe on knowledge diffusion. About 1.7 farmers purchased seeds per village in random no demo villages, and this increased by around 0.67 farmers (40 percent) when adding demonstration plots. The number of farmers purchasing seeds also increases in large and SAO villages. Turning to column 2, the degree centrality of entry points and the number growing the third crop are positively correlated with the number of purchasing farmers and absorb some of the selection effects.[32] Columns 3 and 4 show that the pattern remains when considering the share of farmers purchasing, rather than the absolute number.

The noisier estimates on adoption likely result from the heterogeneous returns of the

[30] Figure A5 shows a positive correlation between eigenvector centrality and the simulated probability that the farmer becomes informed from random seeding. The simulated probability is calculated from 1,000 separate simulations where 5 random entry points are initially drawn, information diffuses through the network for 4 periods, and the probability of passing information to a connected peer is 0.25. These parameter values roughly match the mean knowledge rate in our control group of 0.66. The simulated probability is calculated as the share of the 1,000 simulations where the farmer was informed at the end of the 4 periods.

[31] A local NGO visited each village prior to the 2017 rainy season — a year after BD56 had been introduced — and sold the seeds at a 60 percent discount.

[32] Note we cannot test whether unconnected and less central farmers are more likely to purchase in demonstration plot villages because we did not record the identity of the buyer at the time of sale.

technology. Adopters only benefit from using BD56 if they take advantage of the early harvesting to grow another crop, which only 52% of entry-points did. Interventions to improve learning should therefore have heterogeneous effects since adoption is only beneficial for a subset of the population. Results in the next section confirm that the treatments increased seed purchases only in villages where more farmers are expected to grow an additional crop if they plant BD56.

# 4 Results: Impact Experiment

We now provide our main results from the impact evaluation experiment. We start by reviewing the benefits of cultivating the new short duration rice variety relative to the traditional long-duration variety. This analysis shows that adoption was only profitable for farmers who took advantage of BD56's earlier harvest date to plant an additional crop (between the two main cropping cycles). It confirms that the technology is not profitable for all farmers to adopt. We then use machine learning tools developed by Chernozhukov et al. (2018) to identify a set of farmer baseline characteristics that correlate with the propensity to reap these benefits (which we refer to as a predicted benefit index). Bringing this index into the diffusion analysis, we estimate whether the treatment effects on information diffusion vary with the machine learning prediction that the farmer would actually benefit from planting BD56.

## 4.1 Benefits of the new technology

We compare outcomes for the recipients of BD56 (the entry points in the 192 treated villages) to the recipients of the BD51 seeds (similarly selected in the 64 control villages). First, we do not find any evidence of differential adoption rates (Table A6). Table A7 and Figure A6 show that farmers planting BD56 harvested those fields 25 days earlier than farmers sowing BD51 (in late October rather than mid November). Treatment farmers used this additional month between their two rice crops to increase the likelihood of planting a post-Aman (Rabi) crop. On average, BD56 plots were 27.8 percentage points more likely to be sown with a Rabi crop than BD51 plots. Mustard, pulses, and potatoes were the most frequent short-season Rabi crop induced by the treatment. Importantly for knowledge diffusion, we also find that the probability of growing the Rabi crop is 17 and 11 percentage points higher for large and SAO farmers, respectively (Table A8). In other words, large and SAO-selected entry points were more likely to demonstrate the key benefit of the new variety.

While the BD56 treatment led to a sharp increase in cropping intensity, we still observe that 46 percent of the BD56 plots were left fallow in between the two rice crops.[33] In addition, BD56 naturally leads to lower yields given its shorter duration: the yield of BD56 plots was 31 percent lower than that of the longer duration BD51 plots. We also discovered that BD56 fetched a slightly lower market price, which farmers attributed to less familiarity by millers. In combination, profits during the Aman season were lower by 4,576 taka for BD56 plots, or around 44 percent (Table A9).

The average gain in profit from Rabi cultivation is 1518.881 (60 percent) (Table A9 Column 2). Assuming that all these benefits come from the extensive margin of growing the crop, BD56 led to an increase in Rabi profits of 5,241 taka for farmers that complied by growing the additional crop afforded by the treatment. This suggests that the technology was profitable among the subset of farmers who fully complied by planting the Rabi crop, but was not profitable on average since not all farmers capitalized on this main benefit of the technology.[34]

## 4.2 Heterogeneous benefits of the new technology

We use the impact experiment to estimate every farmer's "predicted benefit" from planting BD56. We use the number of crops grown as our benefit measure since it was necessary to plant a third crop for BD56 to be a profitable investment — an action that was not universally taken by all entry-points. While we can also use profits for this exercise, this measure will only partially reflect the benefits of BD56 because it is subject to factors outside of the farmer's control (e.g weather shocks).[35] We apply the methods developed by Chernozhukov et al. (2018), which first requires generating a linear prediction of the ATE conditional on observed covariates from our door-to-door census, denoted as $z_i$. Our sample of treatment and control farmers is first divided into two samples: a "training" sample where we seek to estimate the predicted benefit index, denoted as $s_0(z_i)$, and a "validation" sample where we seek to verify whether this estimate $\hat{s_0}(z_i)$ is a significant determinant of heterogeneity in the diffusion treatment effect. First for the training sample, we estimate separate LASSO regressions

[33]There are a number of explanations including that farmers were not prepared to grow an additional crop when the rice matured much earlier than anticipated (despite being told of the duration when receiving the seeds), an inability to access land with plows when it is surrounded by maturing rice, and lack of access to capital for planting an additional crop.

[34]Shortening the growing season could also benefit farmers in additional ways that we do not consider. For instance, it may allow them to harvest earlier and allocate time to non-farm activities. It may also be beneficial for farmers that have inadequate access to savings and therefore experience seasonal consumption variability.

[35]Nevertheless, we compute the predicted benefit indices using profits for robustness, and find they correlate strongly with our estimates using the number of crops (Figure A7). Next, to account for the fact that farmers endogenously select where to place the new seed variety, we run the ML algorithm across the full set of 3 plots we asked farmers about – which produces similar results.

for the treatment and control groups to pick which of the observed covariates predict the number of crops grown $y$. Next, we run OLS regressions with the covariates selected by the two LASSO procedures, and recover the beta coefficients from these regressions.

We then turn to the validation dataset and calculate the conditional expectation functions, $E(y_i|D_i = 1, z_i)$ and $E(y_i|D_i = 0, z_i)$, applying the beta coefficients estimated with the training dataset. We take the difference between the two, which serves as the predicted benefit index, $\hat{s_0}(z_i)$. The next step is to verify that the predicted benefit index proxies for the actual BD56 treatment effect. We do this in two ways. First, we add an interaction between the treatment indicator and $\hat{s_0}(z_i) - \bar{s}$ in a regression where the dependent variable is the number of crops grown.[36] Second, we estimate separate treatment effects for the four quartiles of the distribution of $\hat{s_0}(z_i)$. Finally, this process is iterated 100 times, delivering 100 separate sample divisions and 100 estimates of the predicted benefit index.

Our results show that the observed covariates predict treatment-effect heterogeneity in the validation sample. Figure 5 shows the 100 estimates of the ATE and the linear heterogeneity term. The heterogeneous effect is almost always larger than zero[37], suggesting that the predicted benefit index $\hat{s_0}(z_i)$ does proxy for the true heterogenous effect of BD56 on the number of crops grown. In other words, farmers with larger values of $\hat{s_0}(z_i)$ appear more likely to increase cropping intensity if adopting short-duration rice. Figure 6 shows the separate treatment effects by quartile of the predicted benefit index. Treatment effects increase from 0.1 to 0.3 with the predicted benefit index and are largest in the top two quartiles of the distribution of $\hat{s_0}(z_i)$.

## 4.3 Effects of diffusion treatments by predicted benefits

Combining our two experiments (diffusion and impact evaluation experiments) provides a unique opportunity to test whether alternative diffusion strategies deliver information to farmers with different levels of predicted benefits. Specifically, we test whether the impact of H2H demonstration plots or the entry-points selection strategies depend on the predicted benefit index. We possess 100 estimates of the predicted benefit index $\hat{s_0}(z_i)$ for each of the farmers in our estimation data. We use the median value of these estimates as the source of heterogeneity.

Table 8 shows the results. Columns 1 and 2 interact the diffusion treatment variables directly with $\hat{s_0}(z_i)$, while columns 3 and 4 use an indicator for observations with above-

[36] The coefficient on the treatment indicator in this regression measures the average treatment effect, while the coefficient on the interaction between treatment and $\hat{s_0}(z_i) - \bar{s}$ measures whether the predicted benefit index predicts actual treatment-effect heterogeneity.

[37] The minimum value of the predicted benefit index is only just below zero, and on average is equal to 0.24.

median values of this predicted benefit index. The estimates are noisy, but the predicted benefit index is positively associated with hearing about BD56 and having conversations, indicating that information is more likely to flow to farmers with higher returns when entry points are selected randomly and there are no demonstrations. For instance, farmers with an above-median value are 10.1 percentage points more likely to learn about BD56 in random and no demo villages (column 3). From column 4, these same farmers are expected to have .26 more conversations (about 31 percent). However, the point estimates on the interaction terms between the predicted benefit index and the diffusion treatment indicators are generally negative and the coefficients on the five treatment indicators are positive and of similar magnitudes. These findings indicate that while our treatments increased knowledge by either exploiting existing network structure or triggering conversations, the gains in knowledge appear to be concentrated amongst farmers that would be less likely to capitalize on the main benefit of BD56 if adopting.

This finding sheds light on who is induced to have conversations when policymakers intervene with either an alternative seeding strategy or H2H demonstration plots. The evidence suggests that farmers with the lowest probability of reaping the technology's benefits are the most likely to hear about the new technology. One reasonable interpretation is that conversations take place and information is obtained endogenously. Therefore, farmers with the highest returns from obtaining information are more likely to seek out about the new technology regardless of the dissemination strategy.

Intervening to disseminate information to all farmers still adds value even if it prevents adoption by those with low returns. It is important for all farmers to learn, even those who stand to gain less from adoption. It may even be the case that these farmers learn *not to adopt*. Table A10 shows how the number (Column 1) and share (Column 2) of farmers who purchase BD56 varies with the average predicted benefit index for farmers in that village. The coefficients on the interaction terms of average $\hat{s_0}(z_i)$ with our diffusion treatments are large and positive – indicating that villages with higher predicted benefit indices are more likely to purchase seeds under treatment. If anything, the treatments left adoption unchanged, or decreased adoption in villages with lower predicted benefit. Table A11 pools our treatments for precision, and interacts treatment with an indicator for having an average benefit index above the median. The treatments significantly increased seed purchases in villages where farmers are expected to benefit more from BD56.

# 5 A model that rationalizes the experimental results

This section presents a basic diffusion model that allows for endogenous interaction between unlinked agents and can rationalize our experimental findings. Our goal is not to capture all of the strategic elements of network formation.[38] Instead, we opt for the simplest formulation that captures the tradeoffs introduced by our treatments. The policymaker can introduce information to central entry points to take advantage of the existing network structure. Or, the policymaker can seek to induce more communication beyond existing network links. We show how the two approaches act as substitutes in a way that is consistent with our empirical findings.

## 5.1 Model Environment

We consider a village with $N$ farmers indexed by $i \in \{1, 2, ..., N\}$. The village social network is described by the $N \times N$ adjacency matrix $G$, where $g_{ij} = 1$ indicates that farmers $i$ and $j$ are connected. In terms of our data, this corresponds to the baseline social network module.

The policymaker first seeds information to five entry points, indexed by $j = 1, ...5$. Each entry point is now "informed" and can then spread the information to others, as in the standard information cascade model. We assume, for now, that the network is fixed and farmers can only become informed if they receive information that emanated from an entry point. The parameter $q$ represents network transmission: the exogenous probability that any informed farmer passes information to a connected peer. The probability that a farmer $i$ becomes informed, denoted as $h_i$, is a function of the information-passing probability $q$ and the length of the possible paths between $i$ and each entry point $j$. Intuitively, farmers having the greatest number of shortest paths to entry points become the most likely to be informed (high $h_i$). In contrast, a farmer that has no path to any entry point is not informed ($h_i = 0$).

We build on this standard framework by adding the ability to form new connections. In addition to passively waiting for information to arrive, a farmer can increase the probability of becoming informed by forming a connection with an entry point. Suppose the cost of forming a new connection is $c$ and the probability that an entry point will pass information is $p$. Put differently, $p$ is the probability the entry point directly passes along the information to somebody that approaches him, while $q$ reflects transmission from anyone in the network. A new connection between farmer $i$ and entry point $j$ is denoted as $l_{ij} = 1$ if $i$ chooses to

[38]There are several papers looking at different aspects of how information links are formed. These range from differences between actively sharing information and passively listening (Calvó-Armengol, Martí, and Prat, 2015), the types of initial network structures that allow for efficient information aggregation as societies grow large (Acemoglu, Bimpikis, and Ozdaglar, 2014), and the stigma from seeking information when it might signal low ability or understanding (Chandrasekhar, Golub, and Yang, 2019; Banerjee et al., 2018a).

form the link and 0 otherwise. Overall, the probability of gaining information from one of the two channels, denoted as $\mu_i$, is

$$\mu_i = 1 - (1 - h_i) * \prod_j (1 - p)^{l_{ij}} . \tag{3}$$

Finally, we write $v_i$ as the utility of being informed and normalize the utility of not being informed to 0. This is akin to the benefit index in the machine learning exercise above.

## 5.2 The link-formation decision

The simple problem for the farmer is to decide whether to connect or not with each of the entry points. More formally, the farmer's optimization problem is written as:

$$\max_{l_{ij}} v \left( 1 - (1 - h_i) * \prod_j (1 - p)^{l_{ij}} \right) - \sum_j c * l_{ij}. \tag{4}$$

The problem can be simplified to choosing the number of new contacts with entry points, denoted as $m$, since each entry point adds the same probability of learning $p$. The exact decision rule is that the farmer will link with $m$ entry points if and only if:[39]

$$v(1 - h_i)p(1 - p)^{m-1} > c. \tag{5}$$

The farmer seeks to connect directly to entry points for information if the costs of doing so are low ($c$ is low), or if he is in a poor position to obtain information via diffusion in the network ($h_i$ is low). Increasing the probability that an entry point shares useful information ($p$) increases the likelihood of connecting with one entry point ($m = 1$). If $p$ is sufficiently large, then increasing $p$ causes the marginal benefit of connecting with further entry points to decrease because the farmer is likely to obtain the information from the first entry point newly added to her network.

## 5.3 Consistency between the model and experimental findings

This small modification to a standard diffusion framework predicts treatment effects that are in line with our experimental results. The two mechanisms for being informed — receiving information via the existing social network or communicating with an entry point outside

[39] The lefthand side is the difference in expected utility from m and m-1 links. The righthand side is the marginal cost of making that mth link.

the network — can explain the main impacts as well as heterogeneity across the sample. The discussion that follows links the model to our results.

**Effectiveness of entering with Large/SAO farmers:** By being more central in networks, large and SAO-selected farmers facilitate diffusion. Holding networks fixed, the probability $h_i$ in equation (5) increases, causing the likelihood of receiving information to increase. This prediction is the obvious one that supports network-based approaches to identifying entry points when network structure remains fixed. The main effects of large farmer and SAO-based selection on knowledge, and the sensitivity of these effects to conditioning on centrality of entry points (Tables 2 and 6), are both consistent with this standard mechanism.

**Effectiveness of demonstration plots:** We argue that demonstration plots convey active experimentation and signal a farmer that is paying attention to how the new technology performs against its relative alternative. This is nontrivial as Hanna, Mullainathan, and Schwartzstein (2014) show that inattention bias can hinder what farmers could learn from experimentation. As a result, fellow villagers perceive that the demonstrating farmer is more likely to pass useful information, i.e. the parameter $p$ increases with the introduction of demonstration plots. The marginal benefit of forming a link with a single entry point then increases. Correspondingly, the new information link increases the likelihood of becoming informed. The results on knowledge and conversations in Table 2 are consistent with this reasoning.

**Interactions between network-based selection and demonstration plots:** Demonstration plots and seeding with more central farmers are substitutes. Returning to equation (5), seeding with more central farmers increases the probability of learning via existing network links ($h_i$) and therefore reduces the marginal benefit of endogenously seeking information from entry points. The empirical findings show this exactly: demonstration plots have no effect when entry points are more central.

**Network effects:** Farmers directly connected to entry points gain less from demonstration plots because the existing network connection offers a high likelihood of learning through the information-diffusion mechanism. We found that baseline connections with entry points do increase knowledge, but that demonstration plots eliminate this advantage by giving a channel for unconnected farmers to learn. This is compatible with equation (5) where the benefit of making new links with entry points declines with $h_i$.

In sum, our experimental findings, along with the simple model, emphasize two mechanisms that serve to boost learning in networks: optimizing the selection of entry points or inducing communication to facilitate learning. The latter mechanism has received less attention in the literature. However, our results suggest that it gives the policymaker an important alternative for making information diffuse more broadly.

# 6 Concluding Remarks

We have provided experimental evidence on a new mechanism for spreading information about technology in a community of farmers. H2H demonstration plots — where selected farmers cultivate a new technology side by side with an existing one — improve knowledge dissemination relative to a control where new technology is demonstrated on its own. Demonstration plots signal to farmers that an experiment is taking place. They are not a new concept in rural areas: agricultural extension agents often set up large-scale demonstration plots and invite farmers from the region to learn from them. We build on this approach by encouraging a selected group of farmers to set up smaller demonstration plots within their own villages. We find that this method made an additional seven percent of farmers informed about the existence of a new seed variety. H2H demonstration plots worked only for farmers that lacked social connections with adopters. In addition, H2H demonstration plots were most effective for farmers that were most isolated in the network — in terms of their eigenvector centralities. All of these results, when taken together, are consistent with a model where demonstration plots facilitate knowledge transmission by triggering new interactions and conversations among farmers that may not belong to the same social network.

The experiment benchmarked the demonstration plots against policy-relevant alternatives where entry points were selected for their high centralities in the social network. These consisted of entering with the largest farmers and those hand picked by government extension agents. Indeed, these methods do increase the spread of information compared to random selection. Importantly, the gains in knowledge are found to be about the same as the gains from demonstration plots. And in contrast to triggering communication between people without links, the selection treatments seem to be effective because of how they exploit the structure of the baseline social network and because the more central entry points did a better job of demonstrating a key benefit of new technology.

We also showed evidence on who becomes informed through the interventions. The seed variety we introduced has heterogeneous benefits. Specifically, only some farmers could take advantage of the early maturation by increasing cropping intensity. Applying machine-learning methods to identify the characteristics associated with taking this action, we found that farmers who are the most likely to grow an additional crop when adopting BD56 are not those learning from demonstration plots. We found the same result with improved seeding strategies. This finding suggests that intervening to increase information transmission may be ineffective for the highest return individuals who learn even in the absence of these efforts.

At the same time, it is important for all people to become informed, even if they have lower returns and possibly decide not to adopt.

Overall, our analysis highlights the potential for alternate mechanisms of knowledge transmission, beyond those that rely on information cascades through social networks — as typically used by the agricultural extension service in our context. There have been few studies that compare information cascades with alternative methods of spreading knowledge. Focusing on policy, it is important to consider such alternatives because policymakers may face difficulty in identifying the entry points that are theoretically positioned for the best spread of information. Either it could be prohibitively expensive to do a full network survey, or there may not be observable characteristics (such as farm size) that correlate strongly with less observable measures of network centrality. Our results show that improving knowledge can be achieved in these contexts by taking small steps to alert people to the fact that there is something to be learned from informed members in the community, even if they are not from their social network.

# References


Acemoglu, Daron, Kostas Bimpikis, and Asuman Ozdaglar. 2014. "Dynamics of information exchange in endogenous social networks." *Theoretical Economics* 9 (1):41–97.

Akbarpour, Mohammad, Suraj Malladi, and Amin Saberi. 2018. "Just a Few Seeds More: Value of Network Information for Diffusion." *Unpublished* .

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

Banerjee, Abhijit, Emily Breza, Arun Chandrasekhar, and Benjamin Golub. 2018a. "When Less is More: Experimental Evidence on Information Delivery During India's Demonitization." *Unpublished* .

Banerjee, Abhijit, Emily Breza, Arun G Chandrasekhar, Esther Duflo, Matthew O Jackson, and Cynthia Kinnan. 2018b. "Changes in Social Network Structure in Response to Exposure to Formal Credit Markets." *Unpublished* .

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. 2013. "The diffusion of microfinance." *Science* 341 (6144):1236498.

———. 2019. "Using gossips to spread information: Theory and evidence from two randomized controlled trials." *The Review of Economic Studies* 86 (6):2453–2490.

Beaman, Lori, Ariel BenYishay, Mushfiq Mobarak, and Jeremy Magruder. 2018. "Can Network Theory based Targeting Increase Technology Adoption?" *Unpublished* .

Beaman, Lori and Andrew Dillon. 2018. "Diffusion of agricultural information within social networks: Evidence on gender inequalities from Mali." *Journal of Development Economics* 133:147–161.

BenYishay, Ariel and A Mushfiq Mobarak. 2019. "Social learning and incentives for experimentation and communication." *The Review of Economic Studies* 86 (3):976–1009.

Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman. 2014. "Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions." *Econometrica* 82 (4):1273–1301.

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

Calvó-Armengol, Antoni, Joan Martí, and Andrea Prat. 2015. "Communication and influence." *Theoretical Economics* 10 (2):649–690.

Centola, Damon. 2010. "The spread of behavior in an online social network experiment." *Science* 329 (5996):1194–1197.

Chandrasekhar, Arun and Randall Lewis. 2016. "Econometrics of sampled networks." *Unpublished* .

Chandrasekhar, Arun G, Benjamin Golub, and He Yang. 2019. "Signaling, Shame, and Silence in Social Learning." Tech. rep., National Bureau of Economic Research.

Chandrasekhar, Arun G, Horacio Larreguy, and Juan Pablo Xandri. 2020. "Testing models of social learning on networks: Evidence from two experiments." *Econometrica* 88 (1):1–32.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Ivan Fernandez-Val. 2018. "Generic machine learning inference on heterogenous treatment effects in randomized experiments." Tech. rep., National Bureau of Economic Research.

Cole, Shawn, Xavier Giné, and James Vickery. 2017. "How does risk management influence production decisions? Evidence from a field experiment." *The Review of Financial Studies* 30 (6):1935–1970.

Conley, Timothy G and Christopher R Udry. 2010. "Learning about a new technology: Pineapple in Ghana." *American Economic Review* :35–69.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review* 101:2350–2390.

Emerick, Kyle, Alain de Janvry, Elisabeth Sadoulet, and Manzoor H Dar. 2016. "Technological Innovations, Downside Risk, and the Modernization of Agriculture." *American Economic Review* 106 (6):1537–1561.

Foster, Andrew D and Mark R Rosenzweig. 1995. "Learning by doing and learning from others: Human capital and technical change in agriculture." *Journal of Political Economy* 103 (6):1176–1209.

Golub, Benjamin and Matthew O Jackson. 2012. "How Homophily Affects the Speed of Learning and Best-Response Dynamics." *Quarterly Journal of Economics* 127 (3):1287–1338.

Guiteras, Raymond, James Levinsohn, and Ahmed Mushfiq Mobarak. 2015. "Encouraging sanitation investment in the developing world: a cluster-randomized trial." *Science* 348 (6237):903–906.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2014. "Learning through noticing: Theory and evidence from a field experiment." *The Quarterly Journal of Economics* 129 (3):1311–1353.

Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry. 2014. "Agricultural decisions after relaxing credit and risk constraints." *The Quarterly Journal of Economics* 129 (2):597–652.

Kim, David A, Alison R Hwong, Derek Stafford, D Alex Hughes, A James O'Malley, James H Fowler, and Nicholas A Christakis. 2015. "Social network targeting to maximise population behaviour change: a cluster randomised controlled trial." *The Lancet* 386 (9989):145–153.

Kondylis, Florence, Valerie Mueller, and Jessica Zhu. 2017. "Seeing is believing? Evidence from an extension network experiment." *Journal of Development Economics* 125:1–20.

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

Mobius, Markus, Tuan Phan, and Adam Szeidl. 2015. "Treasure hunt: Social learning in the field." Tech. rep., National Bureau of Economic Research.

Munshi, Kaivan. 2004. "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution." *Journal of Development Economics* 73 (1):185–213.

Rigol, Natalia, Reshmaan Hussam, and Benjamin Roth. 2017. "Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design In The Field." *Unpublished* .

Sacerdote, Bruce. 2001. "Peer effects with random assignment: Results for Dartmouth roommates." *The Quarterly Journal of Economics* 116 (2):681–704.

Suri, Tavneet. 2011. "Selection and comparative advantage in technology adoption." *Econometrica* 79 (1):159–209.

Valente, Thomas W. 2012. "Network interventions." *Science* 337 (6090):49–53.

# Tables

Table 1: Differences in baseline characteristics for different entry points

| | Coefficients and SE: | | | |
|---|---|---|---|---|
| | (1) Constant | (2) SAO | (3) Large farmers | (4) p-value (2)-(3) |
| *Network Variables:* | | | | |
| Degree | 4.562*** | 3.582*** | 4.481*** | 0.473 |
| | (0.355) | (1.042) | (0.853) | |
| Eigenvector centrality | 0.089*** | 0.042*** | 0.071*** | 0.030 |
| | (0.006) | (0.012) | (0.011) | |
| Diffusion centrality | -0.010 | 0.643*** | 0.872*** | 0.157 |
| | (0.054) | (0.141) | (0.110) | |
| Betweenness centrality | 164.186*** | 394.084*** | 315.640*** | 0.509 |
| | (27.926) | (103.540) | (69.762) | |
| *Household Characteristics:* | | | | |
| Area cultivated all seasons (bigah) | 9.013*** | 5.865*** | 21.396*** | 0.000 |
| | (0.658) | (1.368) | (2.689) | |
| Times named best farmer | 0.790*** | 4.477*** | 5.589*** | 0.275 |
| | (0.206) | (0.785) | (0.707) | |
| Log revenue per bigah | 10.061*** | -0.016 | -0.014 | 0.970 |
| | (0.057) | (0.077) | (0.075) | |
| Number livestock owned | 3.950*** | -0.008 | 1.968*** | 0.000 |
| | (0.217) | (0.284) | (0.512) | |
| Number of overseas migrants | 0.138*** | -0.021 | -0.026 | 0.881 |
| | (0.031) | (0.039) | (0.037) | |
| Education | 4.647*** | 1.247*** | 0.925* | 0.536 |
| | (0.304) | (0.464) | (0.488) | |
| Age | 42.222*** | 0.712 | 3.594*** | 0.007 |
| | (0.739) | (1.026) | (1.078) | |
| Tubewell owner | 0.097*** | 0.094*** | 0.181*** | 0.108 |
| | (0.022) | (0.036) | (0.051) | |

The data are limited to the 960 selected entry points in the 192 BD56 villages. Each row is the result from a separate regression where the characteristic is regressed on a constant and indicators for SAO and large farmer villages. The omitted group is the villages where demonstrators were selected randomly (meaning the first column is the mean value for random entry points). The standard errors in each regression are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2: Treatment effects on knowledge

| | (1)<br>Heard About | (2)<br>Conversations | (3)<br>Conversations w/ Entry |
|---|---|---|---|
| Random w/ demo | 0.072* | 0.116 | 0.106 |
| | (0.041) | (0.086) | (0.088) |
| SAO no demo | 0.067* | 0.046 | 0.042 |
| | (0.039) | (0.067) | (0.071) |
| SAO w/ demo | 0.065 | 0.074 | 0.096 |
| | (0.040) | (0.089) | (0.086) |
| Large no demo | 0.074** | 0.123* | 0.114* |
| | (0.036) | (0.068) | (0.066) |
| Large w/ demo | 0.049 | 0.108 | 0.113 |
| | (0.044) | (0.075) | (0.079) |
| Strata fixed effects | Yes | Yes | Yes |
| Mean in Random No Demo | 0.60 | 0.84 | 0.72 |
| Number of Observations | 1919 | 1920 | 1920 |
| R squared | 0.171 | 0.212 | 0.250 |

The data are for the 10 random farmers per village that were selected for the information survey. The dependent variable in column 1 is an indicator for having knowledge of BD56. The dependent variable in column 2 is the number of conversations the farmer had with 15 other farmers about BD56 (the five entry points and 10 randomly selected farmers). The dependent variable in column 3 is the number of conversations specifically with entry points. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Note, the estimates on SAO w/ demo, Large w/demo, and Random w/demo are statistically indistinguishable at the 10% level.

Table 3: Peer effects on knowledge, separate for villages with and without demonstration plots

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Peer connections w/ entry points | 0.077** | 0.135*** | | |
| | (0.034) | (0.046) | | |
| Peer connections w/ entry points * Demonstration Village | | -0.116* | | |
| | | (0.061) | | |
| Connected to at least 1 entry point | | | 0.092* | 0.156*** |
| | | | (0.053) | (0.059) |
| Connected to at least 1 entry point * Demonstration Village | | | | -0.130 |
| | | | | (0.097) |
| Total peer connections | -0.004 | -0.005 | -0.002 | 0.002 |
| | (0.003) | (0.006) | (0.003) | (0.005) |
| Total peer connections * Demonstration Village | | 0.000 | | -0.005 |
| | | (0.006) | | (0.005) |
| Demonstration Village | | 0.113** | | 0.132*** |
| | | (0.045) | | (0.045) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean in Random No Demo | 0.60 | 0.60 | 0.60 | 0.60 |
| Number of Observations | 635 | 635 | 635 | 635 |
| R squared | 0.186 | 0.199 | 0.185 | 0.197 |

The data are for the 10 random farmers per village that were selected for the information survey. The data are limited to the 64 villages where entry points where chosen randomly and peer effects can therefore be causally identified. The dependent variable in all regressions is an indicator for having heard of BD56 amongst the 10 randomly surveyed farmers per village The variable *Peer connections w/ entry points* is the number of entry points (from 0 to 5) that the farmer is connected with while *Connected to at least 1 entry point* is an indicator variable for being connected to at least one of the entry points. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 4: Peer effects on the number of conversations with entry points, separate for villages with and without demonstration plots

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Peer connections w/ entry points | 0.050 | 0.179* | | |
| | (0.064) | (0.105) | | |
| Peer connections w/ entry points * Demonstration Village | | -0.261* | | |
| | | (0.137) | | |
| Connected to at least 1 entry point | | | 0.093 | 0.284* |
| | | | (0.098) | (0.147) |
| Connected to at least 1 entry point * Demonstration Village | | | | -0.383* |
| | | | | (0.215) |
| Total peer connections | 0.001 | -0.001 | 0.001 | 0.005 |
| | (0.005) | (0.011) | (0.004) | (0.009) |
| Total peer connections * Demonstration Village | | 0.003 | | -0.005 |
| | | (0.011) | | (0.009) |
| Demonstration Village | | 0.176** | | 0.220** |
| | | (0.087) | | (0.087) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean of Dep Variable | 0.72 | 0.72 | 0.72 | 0.72 |
| Number of Observations | 636 | 636 | 636 | 636 |
| R squared | 0.301 | 0.310 | 0.301 | 0.312 |

The dependent variable in all regressions is the number of entry points that the respondent spoke to about BD56. The data are limited to the 64 villages where entry points where chosen randomly and peer effects can therefore be causally identified. The variable *Peer connections w/ entry points* is the number of entry points (from 0 to 5) that the farmer is connected with while *Connected to at least 1 entry point* is an indicator variable for being connected to at least one of the entry points. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 5: Effects of demonstration plots on knowledge as a function of baseline network centrality

| | (1) | (2) |
|---|---|---|
| Demonstration Village | 0.072* (0.041) | 0.179*** (0.057) |
| Eigenvector Centrality * Demo | | -1.067*** (0.381) |
| Eigenvector Centrality | | 0.935*** (0.325) |
| Strata fixed effects | Yes | Yes |
| Mean in Random No Demo | 0.60 | 0.61 |
| Number of Observations | 639 | 517 |
| R squared | 0.183 | 0.202 |

The dependent variable in both regressions is an indicator for having heard of BD56. The data are limited to the 64 villages where entry points were chosen randomly. *Eigenvector Centrality* is the baseline network centrality of the respondent. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 6: Effects of entry-point treatments on knowledge when conditioning on observable attributes of entry points

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| SAO no demo | 0.068* | 0.047 | 0.056 | 0.033 |
| | (0.038) | (0.039) | (0.039) | (0.040) |
| Large no demo | 0.076** | 0.043 | 0.064* | 0.028 |
| | (0.035) | (0.035) | (0.035) | (0.036) |
| Average degree of entry points | | 0.006*** | | 0.006*** |
| | | (0.002) | | (0.002) |
| Number entry points growing rabi crop | | | 0.020* | 0.022* |
| | | | (0.011) | (0.011) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean in Random No Demo | 0.60 | 0.60 | 0.60 | 0.60 |
| Number of Observations | 960 | 960 | 960 | 960 |
| R squared | 0.173 | 0.179 | 0.176 | 0.182 |

The data are for the 10 random farmers per village that were selected for the information survey and are limited to the 96 villages without demonstration plots. The dependent variable in all regressions is an indicator for having heard of BD56. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 7: Treatment effects on seed purchasing behavior

| | Number of farmers | | Share of village | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Random w/ demo | 0.673 | 0.613 | 0.00622 | 0.00493 |
| | (0.813) | (0.807) | (0.00916) | (0.00917) |
| SAO no demo | 0.697 | 0.392 | 0.0116 | 0.00809 |
| | (0.694) | (0.690) | (0.0107) | (0.0104) |
| SAO w/ demo | 0.116 | -0.154 | 0.00411 | 0.00135 |
| | (0.615) | (0.640) | (0.00744) | (0.00774) |
| Large no demo | 0.272 | -0.171 | 0.00817 | 0.00325 |
| | (0.535) | (0.621) | (0.00782) | (0.00868) |
| Large w/ demo | 0.866 | 0.515 | 0.0195** | 0.0147 |
| | (0.641) | (0.654) | (0.00946) | (0.00914) |
| Average degree of entry points | | 0.0642* | | 0.000567 |
| | | (0.0382) | | (0.000507) |
| Number of entry points growing rabi crop | | 0.166 | | 0.00291 |
| | | (0.156) | | (0.00214) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean in Random No Demo | 1.679 | 1.679 | 0.019 | 0.019 |
| Number of Observations | 168 | 168 | 168 | 168 |
| R squared | 0.085 | 0.106 | 0.091 | 0.106 |

The data are from seed sales that were carried out for each village prior to the 2017 rainy season. We are missing data for 24 of the 192 villages because the seed supply ran out before those villages could be completed. The dependent variables are the number of farmers purchasing BD56 seeds (columns 1-2) and the share of farmers purchasing (columns 3-4). Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

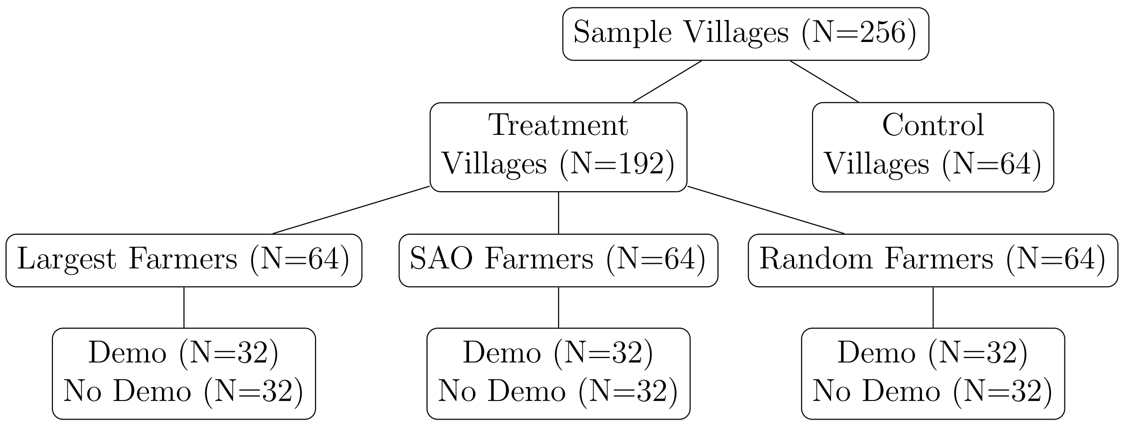
Table 8: Heterogeneous effects on knowledge and conversations by predicted impact of BD56 on number of crops grown

| | Predicted Benefit Heterogeneity | | Index Above Median | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Heard About | Conversation | Heard About | Conversations |
| Random w/ demo | 0.082 | 0.247* | 0.121** | 0.294** |
| | (0.069) | (0.131) | (0.060) | (0.128) |
| SAO no demo | 0.073 | 0.096 | 0.089 | 0.125 |
| | (0.072) | (0.115) | (0.060) | (0.094) |
| SAO w/ demo | 0.096 | 0.184* | 0.111* | 0.253** |
| | (0.064) | (0.110) | (0.057) | (0.099) |
| Large no demo | 0.148** | 0.247** | 0.148*** | 0.284*** |
| | (0.061) | (0.101) | (0.053) | (0.101) |
| Large w/ demo | 0.094 | 0.190 | 0.123* | 0.265*** |
| | (0.083) | (0.116) | (0.068) | (0.096) |
| Heterogeneity | 0.034 | 0.236 | 0.101* | 0.261** |
| | (0.123) | (0.195) | (0.057) | (0.103) |
| Random w/ demo * | -0.020 | -0.430* | -0.078 | -0.318** |
| Heterogeneity | (0.161) | (0.241) | (0.077) | (0.141) |
| SAO no demo * | 0.001 | -0.145 | -0.014 | -0.111 |
| Heterogeneity | (0.164) | (0.282) | (0.072) | (0.115) |
| SAO w/ demo * | -0.097 | -0.369 | -0.072 | -0.337** |
| Heterogeneity | (0.143) | (0.227) | (0.077) | (0.148) |
| Large no demo * | -0.230 | -0.386* | -0.123* | -0.276** |
| Heterogeneity | (0.153) | (0.227) | (0.072) | (0.135) |
| Large w/ demo * | -0.136 | -0.252 | -0.124 | -0.274** |
| Heterogeneity | (0.192) | (0.240) | (0.081) | (0.119) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean in Random No Demo | 0.60 | 0.85 | 0.60 | 0.85 |
| Number of Observations | 1910 | 1911 | 1910 | 1911 |
| R squared | 0.174 | 0.216 | 0.175 | 0.219 |

These regressions test whether the different treatments increase knowledge and spark conversations more (or less) for farmers that are predicted to have the largest impact of BD56 on the number of crops grown. Columns 1 and 2 show linear heterogeneity where the treatment indicators are interacted with $\hat{s_0}(z_i) = E(y_i|D_i = 1, z_i) - E(y_i|D_i = 0, z_i)$ and columns 3 and 4 partition the sample into farmers that are above and below the median in the distribution of $\hat{s_0}(z_i)$. For each farmer we calculate $\hat{s_0}(z_i)$ as the median value across the 100 sample divisions in Figures 5 and 6. The dependent variable in columns 1 and 3 is an indicator for having heard of BD56. The dependent variable in columns 2 and 4 is the number of conversations the farmer had with 15 other farmers about BD56. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

# Figures

Figure 1: Experimental Design




Notes: The 192 BD56 treatment villages were divided into three groups for entry-point selection: random selection, relying on the five largest farmers, and selecting those indicated by the ag. extension officer (SAO). Demonstration plots were set up on half of the 64 villages within each of these arms.

Figure 2: Visualization of the demonstration plot in comparison to the control group

A: Demonstration Plot

B: Control Plot


Notes: Panel A on the left shows an example demonstration plot. The plot on the left side is the BD56 plot while the plot on the right is the popular longer duration variety Swarna. Panel B on the right shows an example from the comparison villages where farmers were only given one marker to denote the BD56 plot.

Figure 3: Timeline of the experiment and data collection

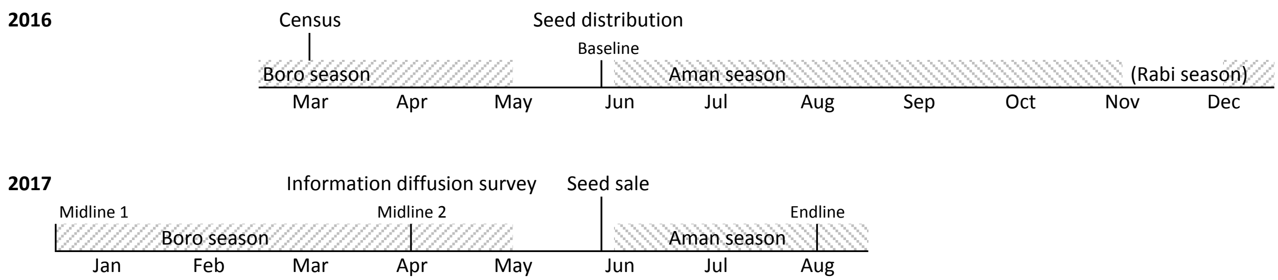
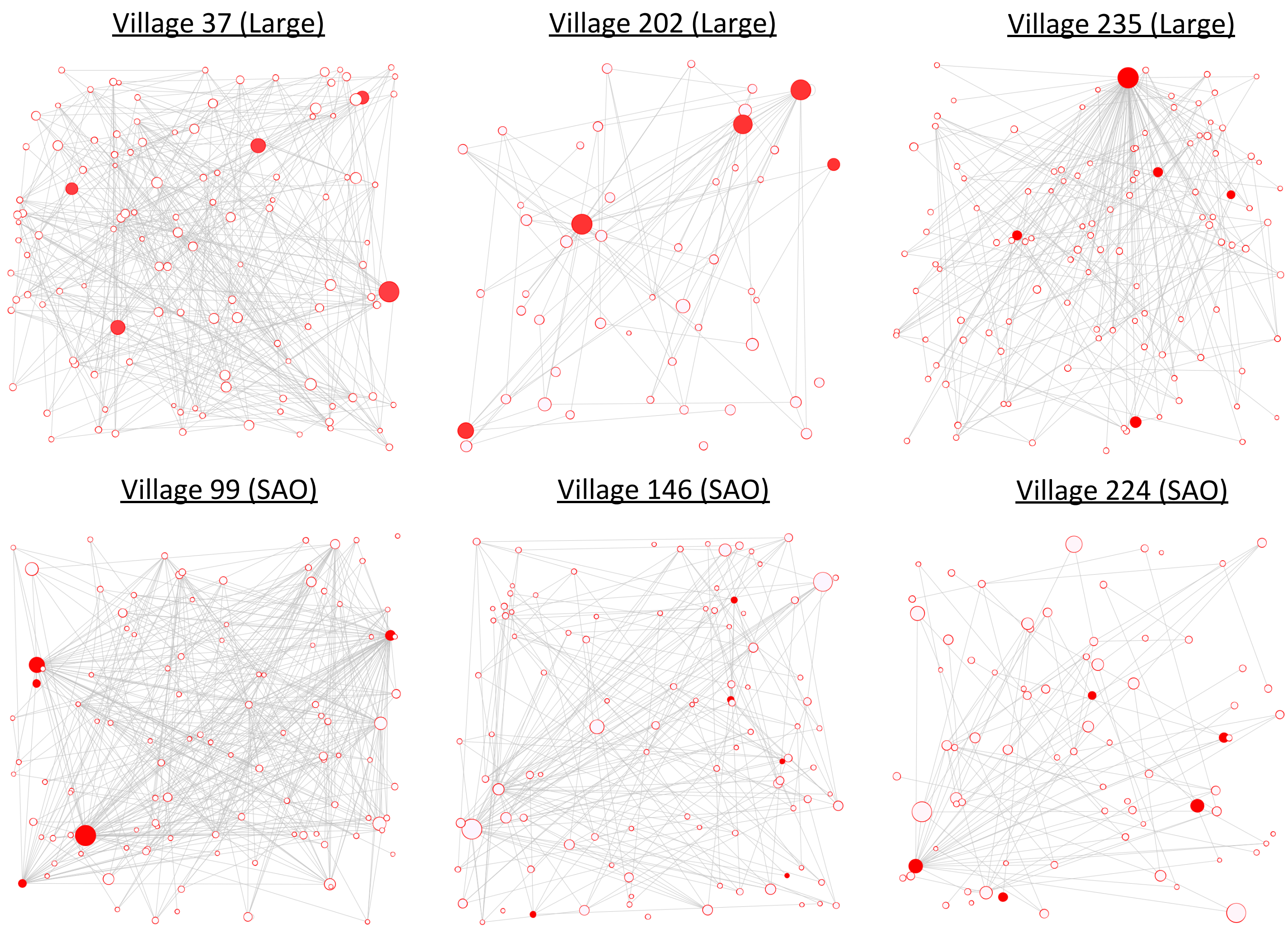
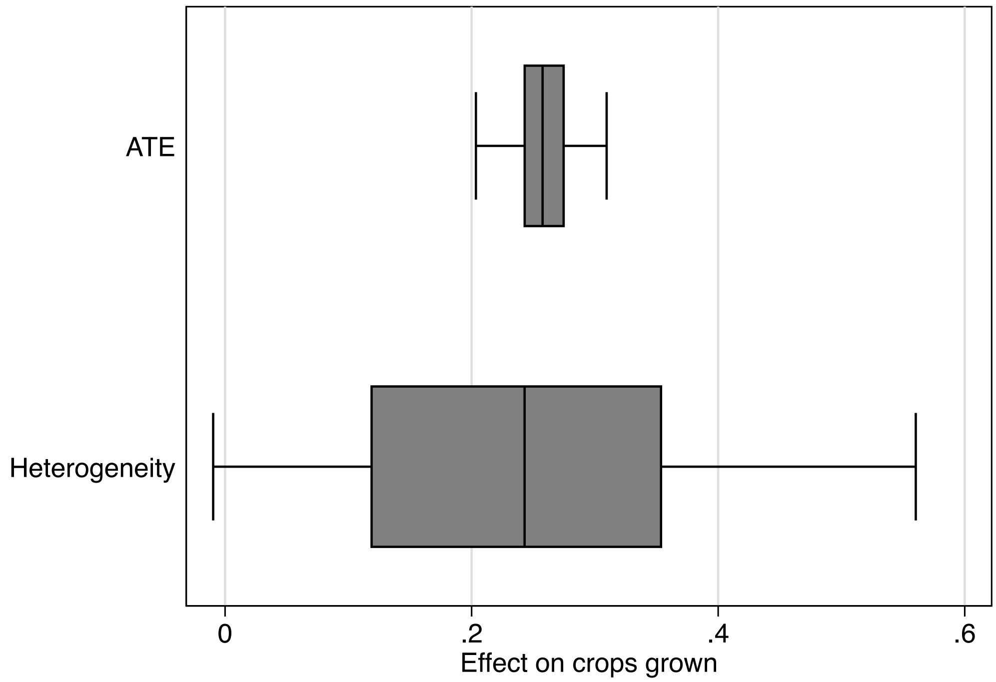


Figure 4: Example network diagrams for 6 villages in the sample



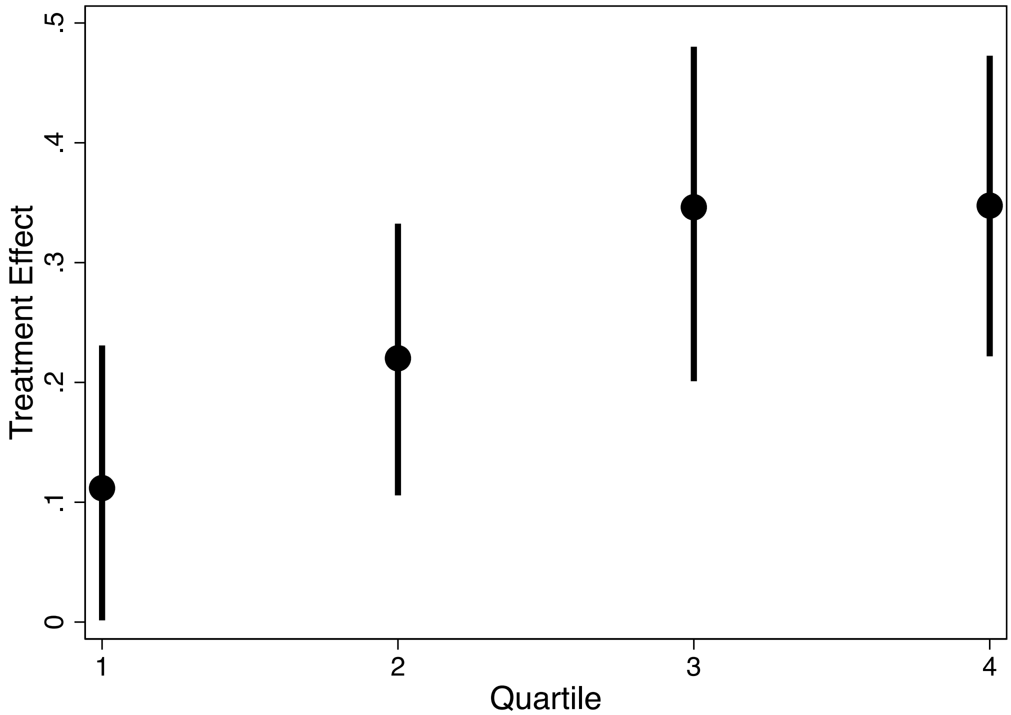

Notes: The figure maps the social network for 6 BD56 villages. The top 3 villages are large-farmer villages and the bottom 3 are villages with SAO selection. The nodes (dots) represent farmers and the size of nodes is proportional to farm size, where larger dots indicate larger farmers. The shaded red dots indicate the 5 farmers chosen as entry points while the hollow red dots denote the remaining farmers.

Figure 5: ATE and heterogeneous effect on number of crops grown




Notes: The figure shows the average treatment effects and the heterogeneous effect on the number of crops grown across 100 equal-sized splits into training and validation datasets. For each split, we estimate separate LASSO regressions for treatment (BD56) and control (BD51) farmers in the training dataset. In each case the number of crops grown is regressed on a set of 24 covariates, $z_i$. Using the selected covariates for each group, we calculate the estimated benefit index for each farmer in the validation dataset as $\hat{s_0}(z_i) = E(y_i|D_i = 1, z_i) - E(y_i|D_i = 0, z_i)$. Using the validation dataset, we then regress the observed number of crops on the treatment, $\hat{s_0}(z_i) - \bar{s_0}$, the interaction between treatment and $\hat{s_0}(z_i) - \bar{s_0}$, and upazila fixed effects. The top bar in the figure shows the distribution of the 100 estimates of the ATE (the coefficients on the treatment indicator). The bottom bar shows the 100 estimates of the heterogeneity effect (the coefficient on the interaction between treatment and $\hat{s_0}(z_i) - \bar{s_0}$). The vertical line represents the average across the 100 splits, the box the interquartile range, and the whiskers give the min and max.

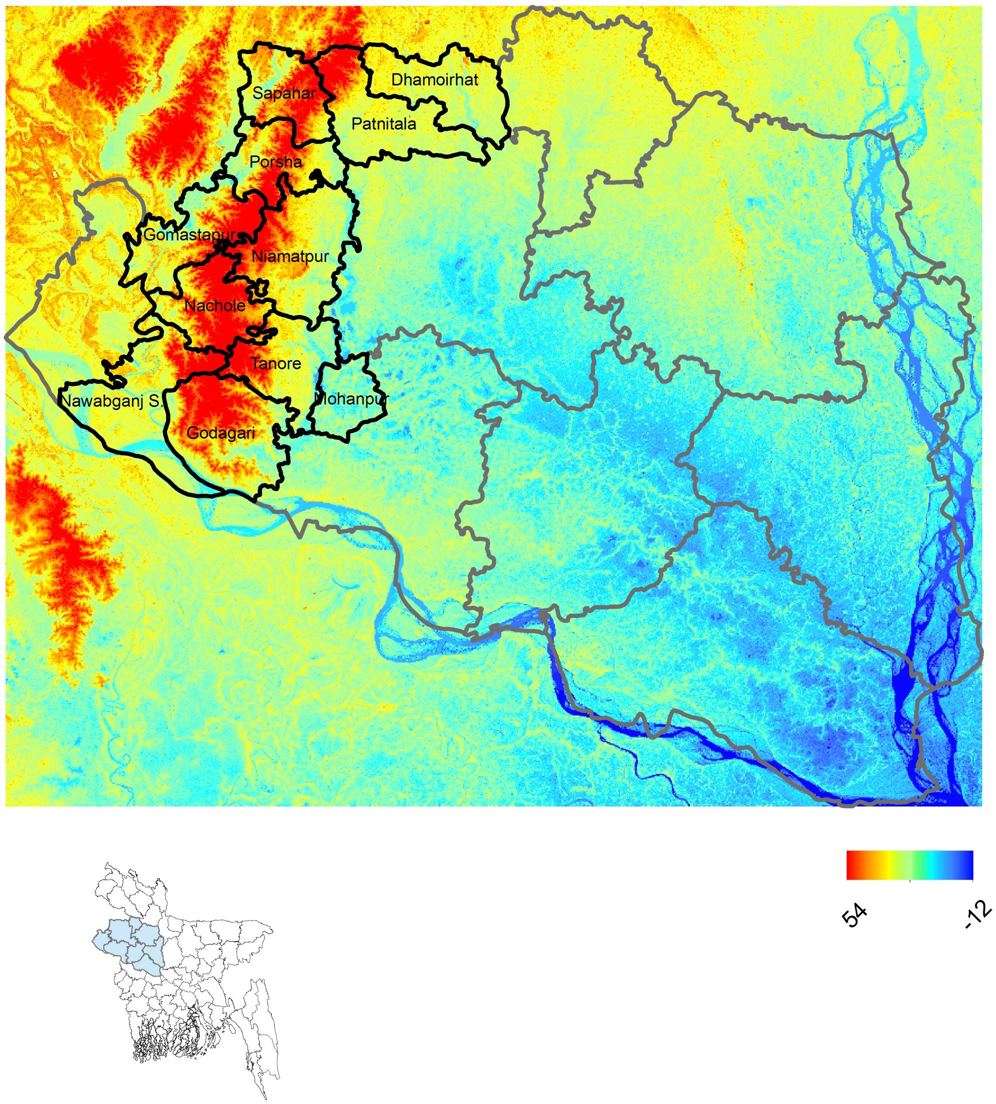
Figure 6: Effects on number of crops grown by quartiles of the predicted effect




Notes: The figure shows the estimated BD56 treatment effects by quartile of the benefit index for 100 equal-sized splits into training and validation datasets. For each split, we estimate separate LASSO regressions for treatment (BD56) and control (BD51) farmers in the training dataset. In each case the number of crops grown is regressed on a set of 24 covariates, $z_i$. Using the selected covariates for each group, we calculate the estimated benefit index for each farmer in the validation dataset as $\hat{s_0}(z_i) = E(y_i|D_i = 1, z_i) - E(y_i|D_i = 0, z_i)$. Using the validation dataset, we then regress the observed number of crops on the treatment and upazila fixed effects *separately for the four quartiles of* $\hat{s_0}(z_i)$. The heavy dots show the averages across the 100 sample divisions while the bands display the range from the 5th to 95th percentiles.

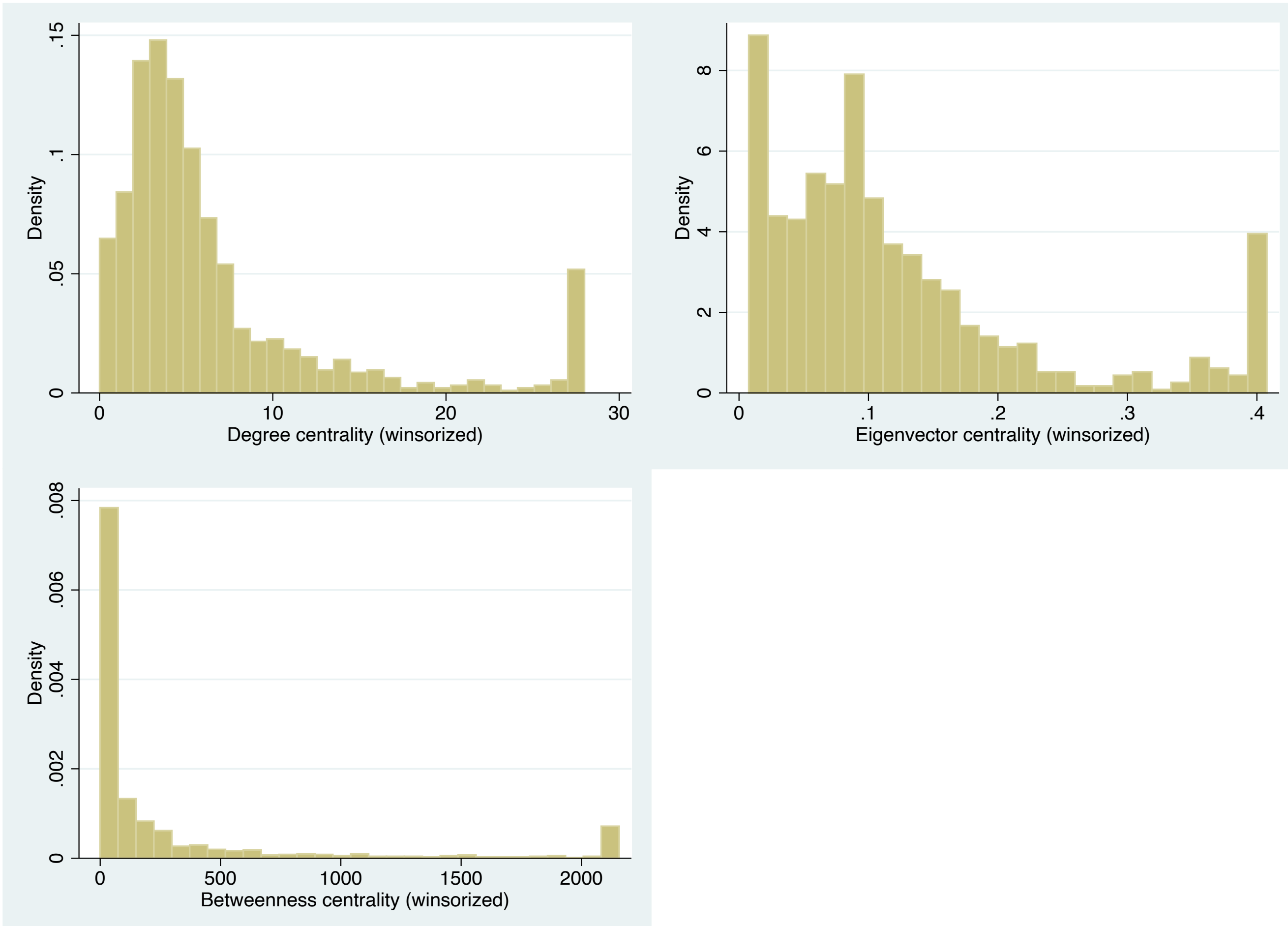
# Appendix: For Online Publication

Figure A1: Location of study area



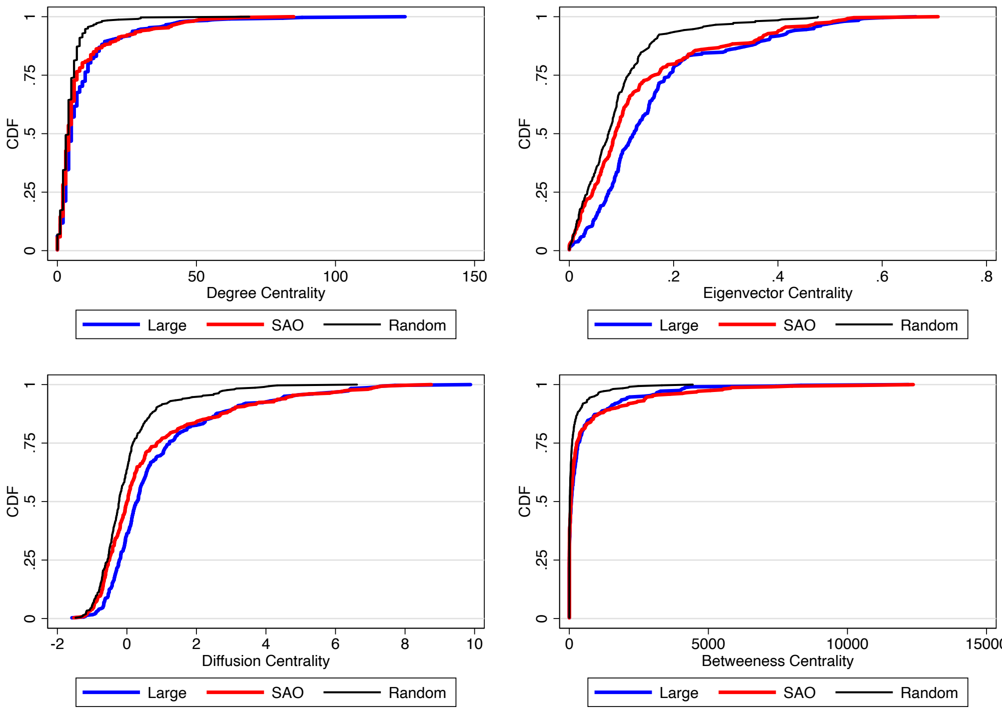

Notes: The figure shows the location of the 11 study Upazilas within Rajshahi district of Bangladesh. The shading corresponds to elevation, measured in meters.

Figure A2: Distribution of centrality measures from social network survey



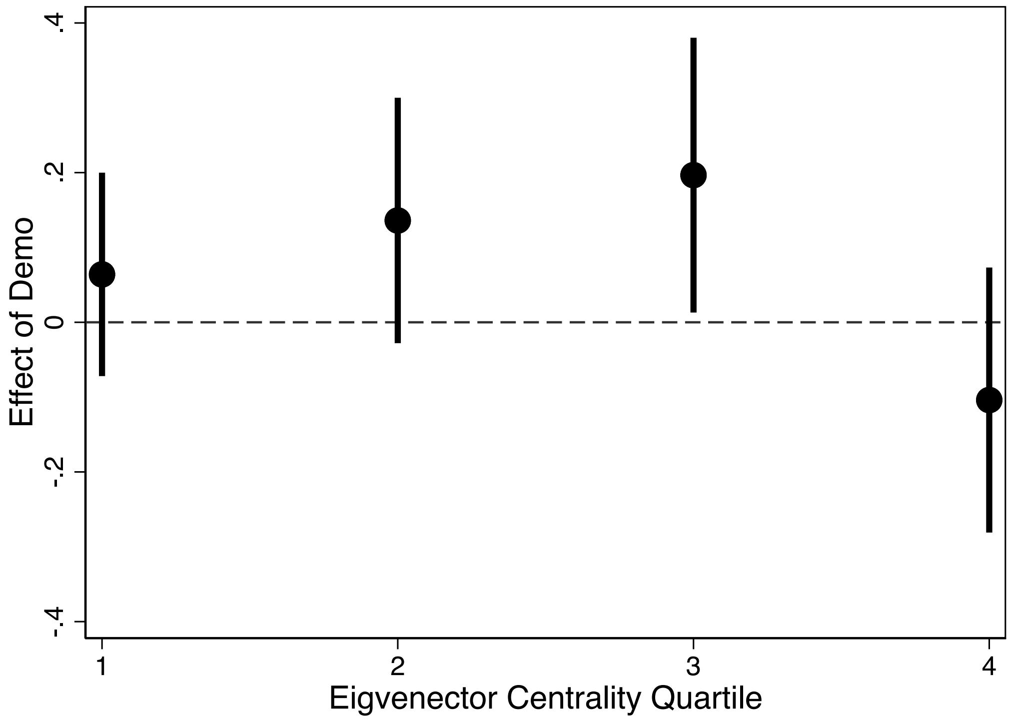

Notes: Figure shows the histograms for the 3 centrality measures from the baseline social network survey with all households (N=21,926).

Figure A3: Cumulative distributions of network statistics for different types of entry points



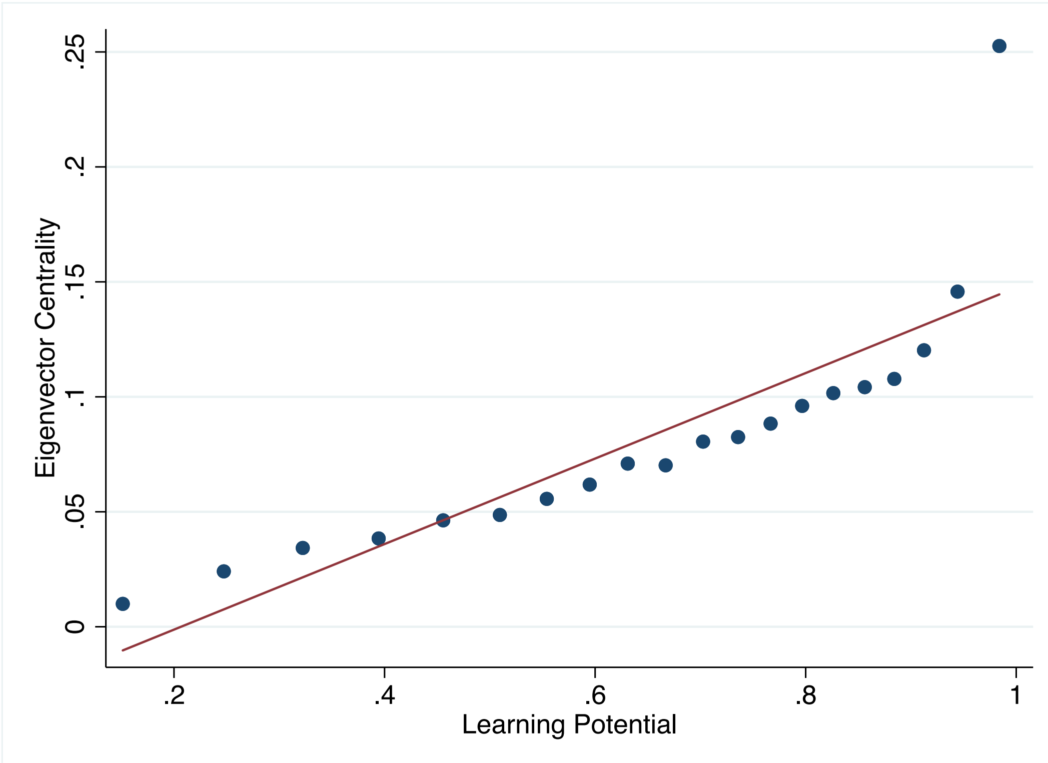

Notes: Each graph shows the cumulative distribution function of the relevant network statistics, separately for the three different types of entry points. The network centrality measures are calculated for each entry point using the baseline social network survey.

Figure A4: Effect of H2H demonstrations by quartile of the eigenvector centrality distribution



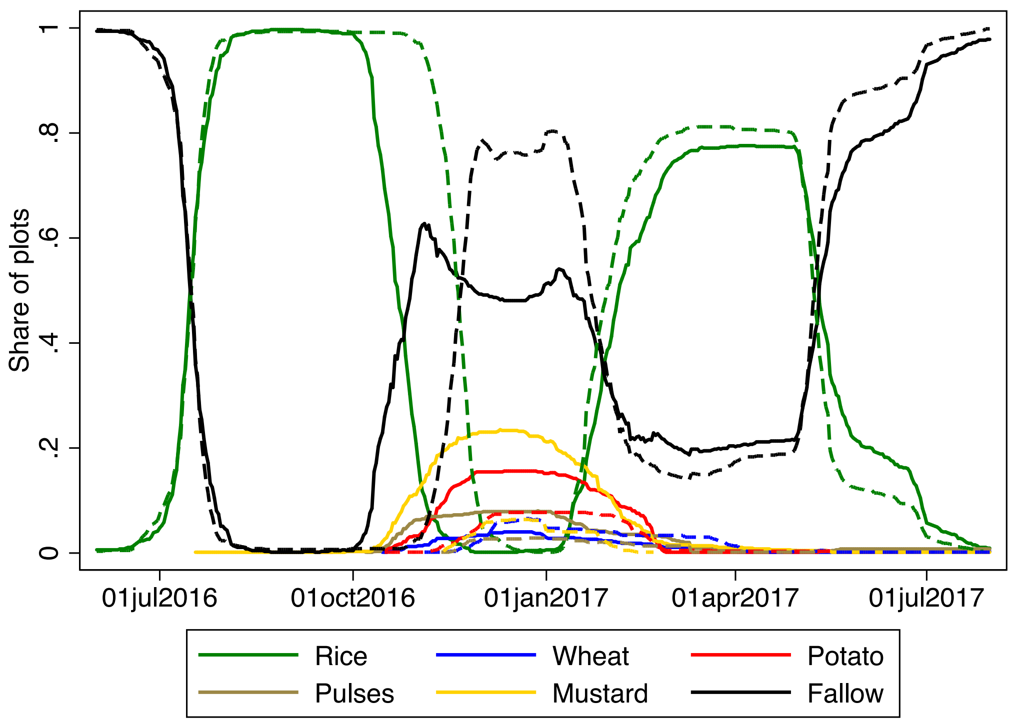

Notes: The data are for the 10 random farmers per village that were selected for the information survey in random villages only (with and without demo). Figure shows the effect of demonstration plots by quartile of the eigenvector centrality distribution for the villages where entry points were selected randomly. The dots represent the effects of the H2H demonstrations, while the heavy lines give 95 percent confidence intervals. The separate effects by quartile are calculated from a single regression where the H2H indicator is interacted with indicators for the different quartiles.

Figure A5: Correlation between eigenvector centrality and network learning potential



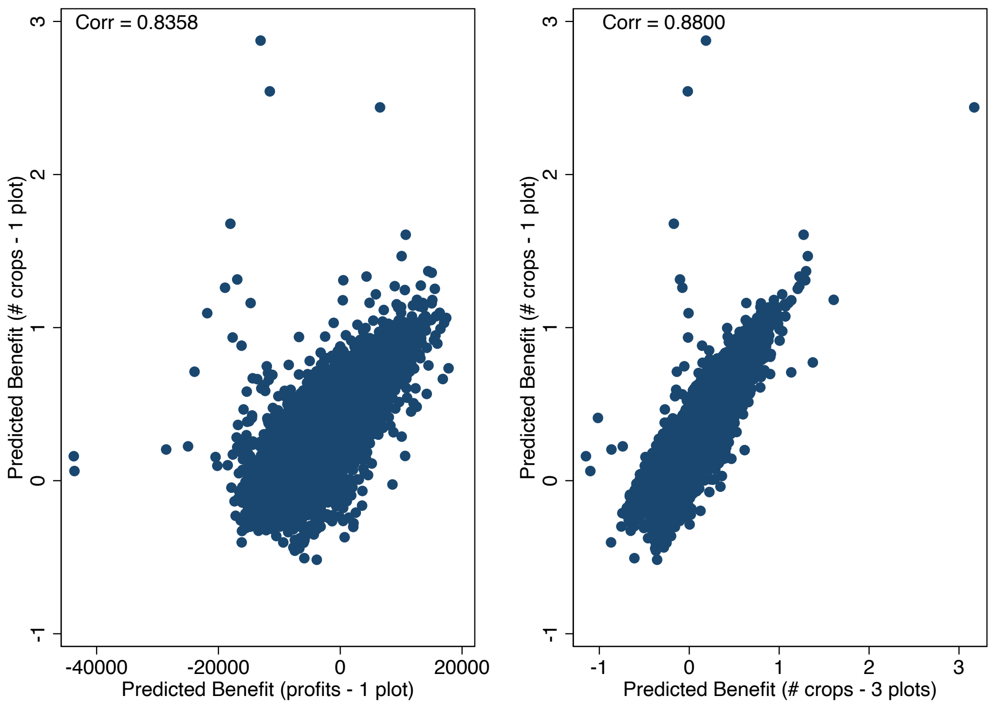

Notes: Figure shows a binned scatterplot of eigenvector centrality against the share of times a farmer becomes informed across 1,000 simulations with random entry points, T=4, and the probability of information being passed along connected nodes of 0.25.

Figure A6: Annual land allocation for plots grown with either BD56 or BD51




Notes: The data are for the plots where either BD56 or BD51 was planted by the entry points. The vertical axis gives the share of plots that were allocated to the crop on the date corresponding to the horizontal axis. The sold lines are for the treatment (BD56) farmers and the dashed lines are for the control (BD51) farmers.

Figure A7: Correlation between predicted benefit indices




Notes: We compute the predicted benefit index using 1) profits as an outcome and 2) all three plots with the number of crops as an outcome. The figure shows the correlation between these predicted benefit indices and our original predicted benefit index (computed using the 'treated' (BD56 or BD51) plot and the number of crops as an outcome).

Table A1: Balance of household characteristics across treatment arms

| | Treatment Arm: | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Control | Large | SAO | Random | Large + Demo | SAO + Demo | Random + Demo | Joint p-value |
| Education | 4.235 | 3.951 | 4.663 | 4.604 | 4.216 | 4.368 | 4.768 | 0.381 |
| | (4.314) | (4.202) | (4.521) | (4.293) | (4.125) | (4.466) | (4.235) | |
| Age | 41.356 | 41.908 | 41.660 | 41.650 | 41.820 | 40.938 | 41.261 | 0.829 |
| | (12.407) | (11.926) | (12.348) | (12.300) | (12.143) | (12.023) | (12.039) | |
| Owns Shallow Tubewell | 0.103 | 0.149 | 0.160 | 0.085 | 0.072 | 0.086 | 0.115 | 0.356 |
| | (0.304) | (0.356) | (0.367) | (0.278) | (0.259) | (0.280) | (0.319) | |
| Aman Rice Area (Bigah) | 4.071 | 4.293 | 5.029 | 4.221 | 4.678 | 4.775 | 4.265 | 0.770 |
| | (5.656) | (5.550) | (12.429) | (6.027) | (5.530) | (5.983) | (5.152) | |
| Aman Other Crop Area (Bigah) | 0.348 | 0.375 | 0.319 | 0.462 | 0.300 | 0.236 | 0.395 | 0.682 |
| | (1.567) | (1.893) | (1.478) | (9.638) | (1.676) | (0.794) | (1.412) | |
| Boro Rice Area (Bigah) | 3.328 | 3.002 | 3.812 | 3.344 | 2.515 | 3.289 | 2.889 | 0.383 |
| | (4.264) | (4.296) | (5.847) | (5.312) | (4.171) | (4.699) | (4.060) | |
| Boro Other Crop Area (Bigah) | 1.125 | 1.252 | 1.332 | 1.140 | 1.478 | 1.100 | 1.346 | 0.840 |
| | (2.454) | (2.513) | (3.325) | (2.362) | (3.043) | (2.219) | (2.358) | |
| Aman Urea Fertilizer (KG per Bigah) | 21.427 | 21.953 | 21.260 | 22.161 | 20.644 | 21.313 | 20.588 | 0.843 |
| | (15.109) | (15.459) | (21.673) | (21.759) | (17.417) | (25.641) | (14.648) | |
| Aman DAP Fertilizer (KG per Bigah) | 15.834 | 16.110 | 15.519 | 16.435 | 14.889 | 15.813 | 15.157 | 0.326 |
| | (11.660) | (15.169) | (20.073) | (13.662) | (6.739) | (15.346) | (10.332) | |
| Aman Rice Yield (KG per Bigah) | 17.756 | 17.499 | 18.063 | 17.989 | 17.477 | 17.570 | 17.837 | 0.927 |
| | (3.814) | (4.180) | (3.263) | (3.089) | (3.280) | (3.851) | (3.483) | |
| Grows Short-Duration Rice | 0.011 | 0.035 | 0.007 | 0.004 | 0.007 | 0.018 | 0.008 | 0.831 |
| | (0.102) | (0.185) | (0.085) | (0.066) | (0.081) | (0.135) | (0.088) | |
| Grows Wheat | 0.236 | 0.260 | 0.243 | 0.190 | 0.372** | 0.231 | 0.286 | 0.353 |
| | (0.425) | (0.439) | (0.429) | (0.393) | (0.483) | (0.422) | (0.452) | |
| Grows Mango | 0.086 | 0.063 | 0.071 | 0.093 | 0.076 | 0.076 | 0.063 | 0.958 |
| | (0.280) | (0.242) | (0.256) | (0.290) | (0.265) | (0.266) | (0.244) | |
| Grows Potato | 0.083 | 0.052 | 0.086 | 0.082 | 0.074 | 0.061 | 0.077 | 0.872 |
| | (0.275) | (0.222) | (0.281) | (0.274) | (0.261) | (0.239) | (0.266) | |
| Grows Pulses | 0.047 | 0.103 | 0.095 | 0.076 | 0.077 | 0.075 | 0.048 | 0.518 |
| | (0.212) | (0.305) | (0.293) | (0.265) | (0.266) | (0.264) | (0.215) | |
| Grows Onion | 0.049 | 0.037 | 0.050 | 0.021* | 0.047 | 0.039 | 0.057 | 0.275 |
| | (0.215) | (0.188) | (0.219) | (0.144) | (0.211) | (0.194) | (0.232) | |
| Grows Garlic | 0.017 | 0.009 | 0.014 | 0.013 | 0.009 | 0.017 | 0.006 | 0.429 |
| | (0.128) | (0.096) | (0.118) | (0.113) | (0.095) | (0.130) | (0.078) | |

The summary statistics are calculated using the door-to-door census with 21,926 households. Each column shows mean values of each variable for either the control group or one of the six treatment groups. Standard deviations are reported in parentheses below each mean value. Asterisks indicate a statistically significant difference (1% ***, 5% **, and 10% * ) between that arm and the control arm, where p-values are calculated by regressing each variable on a constant and indicators for each of the six treatment groups (standard errors adjusted for clustering at the village level). The final column shows the joint p-value of each of these regressions. Aman refers to the wet season prior to the door-to-door baseline (2015) and Boro refers similarly to the most recent dry season (2015-2016). 1 Bigah = 0.33 Acres.

Table A2: Balance of household characteristics for entry points

| | Control (BD51) | BD56 Treatment | p-value |
|---|---|---|---|
| Education | 5.361 | 5.392 | 0.832 |
| | (4.620) | (4.643) | |
| Age | 43.326 | 43.690 | 0.577 |
| | (12.585) | (12.225) | |
| Owns Shallow Tubewell | 0.178 | 0.191 | 0.528 |
| | (0.383) | (0.393) | |
| Aman Rice Area (Bigah) | 8.553 | 8.977 | 0.523 |
| | (11.097) | (10.494) | |
| Aman Other Crop Area (Bigah) | 0.514 | 0.616 | 0.415 |
| | (1.648) | (2.037) | |
| Boro Rice Area (Bigah) | 6.483 | 6.247 | 0.795 |
| | (7.781) | (8.606) | |
| Boro Other Crop Area (Bigah) | 2.063 | 2.299 | 0.454 |
| | (3.709) | (3.951) | |
| Aman Urea Fertilizer (KG per Bigah) | 21.879 | 21.316 | 0.541 |
| | (24.672) | (15.565) | |
| Aman DAP Fertilizer (KG per Bigah) | 16.363 | 15.978 | 0.735 |
| | (17.281) | (18.909) | |
| Aman Rice Yield (KG per Bigah) | 17.888 | 17.566 | 0.207 |
| | (3.797) | (3.881) | |
| Grows only rice | 0.389 | 0.337 | 0.132 |
| | (0.488) | (0.473) | |
| Grows Short-Duration Rice | 0.017 | 0.024 | 0.488 |
| | (0.129) | (0.153) | |
| Grows Wheat | 0.304 | 0.312 | 0.946 |
| | (0.460) | (0.464) | |
| Grows Mango | 0.126 | 0.117 | 0.792 |
| | (0.332) | (0.322) | |
| Grows Potato | 0.121 | 0.098 | 0.359 |
| | (0.327) | (0.297) | |
| Grows Pulses | 0.078 | 0.111 | 0.127 |
| | (0.268) | (0.314) | |
| Grows Onion | 0.048 | 0.068 | 0.323 |
| | (0.215) | (0.252) | |
| Grows Garlic | 0.013 | 0.028 | 0.065 |
| | (0.115) | (0.166) | |

The analysis uses the door-to-door census conducted at the beginning of the experiment. Data are limited to the 1,747 entry points that consented to participate. Each column shows mean values and standard deviations are reported in parentheses below. The final column shows the p-value for the comparison of means, based on a regression of each characteristic on the treatment indicator and Upazila (strata) fixed effects. Standard errors are clustered at the village level.

Table A3: Differences between SAO selected and random farmers, adjusting for farm size

| | Degree | | Eigenvector | | Betweenness | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| SAO-based selection | 3.582*** | 1.917** | 0.042*** | 0.025** | 394.084*** | 281.922*** |
| | (1.044) | (0.838) | (0.012) | (0.011) | (103.649) | (89.698) |
| Farm Size | | 0.284*** | | 0.003*** | | 19.124*** |
| | | (0.053) | | (0.000) | | (4.634) |
| Mean in random group | 4.56 | 4.56 | 0.09 | 0.09 | 164.19 | 164.19 |
| Number of Observations | 639 | 639 | 511 | 511 | 639 | 639 |
| R squared | 0.037 | 0.221 | 0.036 | 0.175 | 0.033 | 0.094 |

The data are limited to the 640 selected entry points in the random and SAO villages. The dependent variables are degree centrality (columns 1-2), eigenvector centrality (columns 3-4), and betweenness centrality (columns 5-6). Farm size is the total sum of cultivated area (across all three agricultural seasons). The omitted group in each regression is the villages where demonstrators were selected randomly. The standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table A4: Effects of hitting a large-farmer entry point

| | (1) All Villages | (2) Random Villages |
|---|---|---|
| At least 1 large entry point | 0.079** | 0.123* |
| | (0.032) | (0.071) |
| At least 1 large entry point * Demonstration Village | -0.101** | -0.131 |
| | (0.048) | (0.104) |
| Demonstration Village | 0.078** | 0.119** |
| | (0.038) | (0.047) |
| Strata fixed effects | Yes | Yes |
| Mean of Dep Variable | 0.67 | 0.64 |
| Number of Observations | 1919 | 639 |
| R squared | 0.171 | 0.189 |

The data are for the 10 random farmers per village that were selected for the information survey. Column 1 is for all 192 BD56 villages, column 2 is for the 64 villages where entry points were selected randomly. *At least 1 large entry point* is an indicator for villages where one of the five largest farmers was selected as an entry point. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table A5: Treatment effects as a function of network centrality

| | Variable Interacted with Treatments: | |
|---|---|---|
| | (1) | (2) |
| | Eigenvector Centrality | Degree |
| Random w/ demo | 0.170*** | 0.118** |
| | (0.059) | (0.046) |
| SAO no demo | 0.131* | 0.092* |
| | (0.072) | (0.051) |
| SAO w/ demo | 0.186*** | 0.058 |
| | (0.063) | (0.047) |
| Large no demo | 0.213*** | 0.102** |
| | (0.063) | (0.047) |
| Large w/ demo | 0.155** | 0.037 |
| | (0.071) | (0.054) |
| Heterogeneity Variable | 0.912*** | 0.005 |
| | (0.308) | (0.005) |
| SAO no demo * Heterogeneity Variable | -0.784 | -0.006 |
| | (0.528) | (0.007) |
| SAO w/ demo * Heterogeneity Variable | -0.964** | 0.003 |
| | (0.470) | (0.007) |
| Large w/ demo * Heterogeneity Variable | -0.967* | 0.003 |
| | (0.573) | (0.008) |
| Large no demo * Heterogeneity Variable | -1.192** | -0.006 |
| | (0.474) | (0.007) |
| Random w/ demo * Heterogeneity Variable | -1.062*** | -0.009* |
| | (0.375) | (0.005) |
| Strata fixed effects | Yes | Yes |
| Mean in Random No Demo | 0.66 | 0.66 |
| Mean of Hetero. Variable | 0.10 | 4.49 |
| Number of Observations | 1463 | 1919 |
| R squared | 0.172 | 0.174 |

The data are for the 10 random farmers per village that were selected for the information survey. Each column shows results of a separate regression where the variable in the column title is being interacted with the 5 treatment variables. The coefficient labeled "Heterogeneity Variable" is the direct effect of the variable in the column title and the 5 interaction terms between that variable and the treatments are at the bottom of the table. The dependent variable in all columns is an indicator for having knowledge of BD56. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table A6: Analysis of take up by entry points

| | (1) | (2) | (3) |
|---|---|---|---|
| Treatment village | -0.025 | | |
| | (0.035) | | |
| Demo Village | | 0.060 | |
| | | (0.039) | |
| Random + Demo | | | 0.085 |
| | | | (0.062) |
| SAO | | | 0.076 |
| | | | (0.066) |
| SAO + Demo | | | 0.067 |
| | | | (0.067) |
| Large | | | 0.053 |
| | | | (0.067) |
| Large + Demo | | | 0.157** |
| | | | (0.068) |
| Strata (Upazila) Fixed Effects | Yes | Yes | Yes |
| Mean in omitted group | 0.71 | 0.65 | 0.69 |
| Number of Observations | 1795 | 953 | 953 |
| R squared | 0.046 | 0.059 | 0.064 |

The data are from the first midline with 1,795 entry points. Column 1 uses all observations and columns 2 and 3 use only observations from treatment (BD56) villages. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table A7: Cultivation practices by treatment

| | Panel A: Plots with either BD56 or BD51 | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Harvest Date | 2nd Crop | Boro Crop | N Crops |
| Treated Village | -25.350*** | 0.278*** | -0.035 | 0.243*** |
| | (1.384) | (0.035) | (0.039) | (0.046) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean in Control | 8.64 | 0.24 | 0.82 | 2.06 |
| Number of Observations | 1242 | 1242 | 1242 | 1242 |
| R squared | 0.381 | 0.284 | 0.257 | 0.278 |

| | Panel B: TOT using 3 random plots per farmer | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Harvest Date | 2nd Crop | Boro Crop | N Crops |
| BD56 plot | -30.494*** | 0.344*** | -0.129 | 0.215* |
| | (3.940) | (0.081) | (0.113) | (0.130) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean in Control | 9.47 | 0.21 | 0.82 | 2.03 |
| Number of Observations | 4174 | 4174 | 4174 | 4174 |
| R squared | 0.325 | 0.301 | 0.230 | 0.286 |

Panel A limits the data to plots where either BD56 or BD51 was planted. Panel B uses the data on 3 randomly selected plots for which we asked production information during all surveys. Panel B shows treatment on the treated (TOT) estimates where the BD56 plot indicator is instrumented with random assignment to the BD56 treatment group. The dependent variable in column 1 is the date of the harvest, measured in days after November 10, 2016. The dependent variables in columns 2 and 3 are indicators for whether the plot was sown with the Rabi (in-between) crop and the Boro (dry-season) crop. The dependent variable in column 4 is the total number of crops grown across all seasons. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table A8: Cultivation practices by treatment and type of entry point

| | (1) Harvest Date | (2) 2nd Crop | (3) Boro Crop | (4) N Crops |
|---|---|---|---|---|
| Treated Village | -25.315*** | 0.175*** | -0.073 | 0.102 |
| | (1.603) | (0.048) | (0.051) | (0.072) |
| Treatment Village * SAO | -0.219 | 0.110* | 0.100* | 0.210*** |
| | (2.518) | (0.061) | (0.056) | (0.079) |
| Treatment Village * Large | 0.533 | 0.173*** | 0.002 | 0.174** |
| | (2.302) | (0.065) | (0.062) | (0.085) |
| SAO | 1.535* | -0.040 | -0.047 | -0.087** |
| | (0.796) | (0.033) | (0.029) | (0.035) |
| Large | 0.784 | -0.051 | -0.014 | -0.065* |
| | (0.801) | (0.031) | (0.029) | (0.034) |
| Mean in Control | 8.64 | 0.24 | 0.82 | 2.06 |
| Number of Observations | 1242 | 1242 | 1242 | 1242 |
| R squared | 0.382 | 0.291 | 0.262 | 0.286 |

The data are limited to the plots where either BD56 or BD51 was planted by the entry points. The dependent variable in column 1 is the date of the harvest, measured in days after November 10, 2016. The dependent variables in columns 2 and 3 are indicators for whether the plot was sown with the Rabi (in-between) crop and the Boro (dry-season) crop. The dependent variable in column 4 is the total number of crops grown across all season. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table A9: Profitability of BD56 and BD51 plots

| | Panel A: Plots with either BD56 or BD51 | | | |
|---|---|---|---|---|
| | (1) Aman | (2) Rabi | (3) Boro | (4) Total |
| Treated Village | -4576.411*** | 1518.881*** | -882.748 | -4161.562*** |
| | (248.751) | (451.867) | (619.309) | (767.897) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean in Control | 10309.28 | 2508.40 | 11900.65 | 24954.19 |
| Number of Observations | 1200 | 1205 | 1228 | 1156 |
| R squared | 0.396 | 0.438 | 0.314 | 0.380 |

| | Panel B: TOT using 3 random plots per farmer | | | |
|---|---|---|---|---|
| | (1) Aman | (2) Rabi | (3) Boro | (4) Total |
| BD56 plot | -4242.751*** | 2512.619* | -3283.277* | -5378.734** |
| | (793.682) | (1386.549) | (1952.855) | (2531.113) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean in Control | 9283.63 | 2682.34 | 11806.87 | 23925.79 |
| Number of Observations | 4111 | 4078 | 4117 | 3972 |
| R squared | 0.219 | 0.408 | 0.274 | 0.364 |

Panel A limits the data to plots where either BD56 or BD51 was planted. Panel B uses the data on 3 randomly selected plots for which we asked production information during all surveys. Panel B shows treatment on the treated (TOT) estimates where the BD56 plot indicator is instrumented with random assignment to the BD56 treatment group. The dependent variables are profits per bigah, measured in Bangladeshi Taka (BDT). Approximately 80 BDT=1USD and 3 bigah = 1 acre. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table A10: Heterogeneous effects on seed purchasing by village-level predicted benefits

| | (1) Number farmers | (2) Share Purchasing |
|---|---|---|
| Random w/ demo | -1.767 | -0.022 |
| | (1.527) | (0.016) |
| SAO no demo | -0.794 | 0.006 |
| | (1.606) | (0.029) |
| SAO w/ demo | -1.834 | -0.023 |
| | (1.287) | (0.014) |
| Large no demo | -1.079 | -0.010 |
| | (1.250) | (0.014) |
| Large w/ demo | 0.148 | 0.002 |
| | (1.472) | (0.018) |
| Benefit Index | -9.951*** | -0.138*** |
| | (3.287) | (0.043) |
| Random w/ demo * Benefit Index | 7.356 | 0.084 |
| | (5.291) | (0.058) |
| SAO no demo * Benefit Index | 3.789 | -0.001 |
| | (5.068) | (0.082) |
| SAO w/ demo * Benefit Index | 5.731 | 0.080* |
| | (3.858) | (0.046) |
| Large no demo * Benefit Index | 3.703 | 0.050 |
| | (3.147) | (0.035) |
| Large w/ demo * Benefit Index | 0.852 | 0.040 |
| | (3.614) | (0.048) |
| Strata fixed effects | Yes | Yes |
| Mean in Random No Demo | 1.68 | 0.02 |
| Number of Observations | 168 | 168 |
| R squared | 0.159 | 0.164 |

The data are from seed sales that were carried out for each village prior to the 2017 rainy season. We are missing data for 24 of the 192 villages because the seed supply ran out before those villages could be completed. The dependent variables are the number of farmers purchasing BD56 seeds (column 1) and the share of farmers purchasing (column 2). The benefit index is the same as in Table 8, but averaged at the village level. Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table A11: Heterogeneous effects on seed purchasing by village-level predicted benefits, all treatments pooled

| | Number of farmers | | Share of village | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Pooled Treatment | -1.121 | -0.268 | -0.010 | -0.002 |
| | (1.121) | (0.798) | (0.013) | (0.010) |
| Benefit Index | -10.030*** | | -0.138*** | |
| | (3.226) | | (0.041) | |
| Pooled Treatment * Benefit Index | 4.516 | | 0.052 | |
| | (2.879) | | (0.034) | |
| Above Median Village Benefit Index | | -1.462 | | -0.019 |
| | | (0.948) | | (0.012) |
| Pooled Treatment * Above Median | | 1.279 | | 0.019 |
| | | (0.969) | | (0.012) |
| Strata fixed effects | Yes | Yes | Yes | Yes |
| Mean in Random No Demo | 1.68 | 1.68 | 0.02 | 0.02 |
| P-value Treat+ Index*Treat | 0.076 | 0.071 | 0.061 | 0.012 |
| Number of Observations | 168 | 168 | 168 | 168 |
| R squared | 0.129 | 0.086 | 0.128 | 0.083 |

The data are from seed sales that were carried out for each village prior to the 2017 rainy season. We are missing data for 24 of the 192 villages because the seed supply ran out before those villages could be completed. The dependent variables are the number of farmers purchasing BD56 seeds (columns 1-2) and the share of farmers purchasing (columns 3-4). The pooled treatment variable is a binary indicator for being in one of the five treatment arms (Random w/ demo, SAO w/ demo, SAO no demo, Large w/ demo, or Large no demo). The benefit index is the same as in Table 8. but averaged at the village level. Columns 1 and 3 use a continuous measure of the benefit index, while columns 2 and 4 use an indicator for villages above the median. Robust standard errors are in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.