Network-Based Targeting with Heterogeneous Agents for Improving Technology Adoption

Aranya Chakraborty

McGill University

Motivation

- Low adoption of modern technologies in developing countries (Foster and Rosenzweig, 2010).
- One of the key reasons: information constraints (Magruder, 2018).
- Social networks can facilitate technology adoption by improving diffusion (Foster and Rosenzweig, 1995).
- What is the most effective way of using existing social ties to improve diffusion?
  - Targeting vs. random seeding. (Akbarpour et al., 2020)
  - For targeting, need to seed agents based on their positions in the network. (Beaman et al., 2021)
  - **Key Assumption:** For the purpose of diffusion, the agents are homogeneous otherwise.

This Study

What if agents differ in terms of benefits from a new technology, with this having *direct* consequences for diffusion?

- How does it affect the optimality of network-based targeting strategies?
- What are the policy concerns and recommendations in such a scenario?

To answer these questions:

- Theoretically model the scenario where agents are learning about heterogeneous benefits from each other.
- Use simulations to characterize the outcomes of different targeting strategies.
- Test predictions using data on the diffusion of pit planting in Malawi.

Contributions

1. Using networks to improve technology adoption
Banerjee et al. (2013, 2019), Beaman et al. (2021)

- Evidence that the success of network-based targeting strategies depend on the population level heterogeneity.

2. Effect of population heterogeneity in social learning
Munshi (2004), Conley and Udry (2010)

- Formalize agents learning from their network about a technology having heterogeneous benefits.

3. Characterizing opinion leaders in diffusing new knowledge
Feder and Savastano (2006), Maertens (2017)

- Based on population heterogeneity, characterize opinion leaders in network-based targeting.

Elements of the model

Households' face two-step adoption decision:

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

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

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

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

▸ Full Model

Elements of the model: Timeline

1. At each t, uninformed household i decide whether or not to get informed.
2. To decide, uninformed households collect information on beliefs (p_{jt-1}) from their peers $j \in \mathcal{I}$, formed in the last period. Household i use DeGroot averaging to calculate $\hat{p}_{it} = \sum_{j \in \mathcal{I}} G_{ij} p_{jt-1}$ (**Note:** $p_i^* = \sum_{j \in \mathcal{I}} G_{ij} p_j^*$).
3. On the basis of $\hat{p}_{it}$, they decide whether or not to become informed.
4. If not informed ($e_{it} = 0$): $p_{it} = \hat{p}_{it}$, and next period repeat from 1. If informed ($e_{it} = 1$): p_i^* is known and adoption decisions are made on the basis of that. $p_{is} = p_i^* \ \forall s \geq t$.

Implications

- Let's simplify: $\Omega = \{\omega_H, \omega_L\}$ and $p^*_{iH} := p^*_i(\omega_H)$.
- In step 2 the household will adopt the new technology iff:

$$p^*_{iH} \geq \frac{c_i + (\pi^T - \pi^N(\omega_L))}{(\pi^N(\omega_H) - \pi^N(\omega_L))} = \bar{p}^*_{iH}.$$

- In step 1 the household i will choose to get informed at time t iff:

$$p^H_{it} \geq \bar{p}^*_{iH} + \frac{\eta_i}{(\pi^N(\omega_H) - \pi^N(\omega_L))} = \bar{p}^*_{iH} + \bar{\eta}_i.$$

- Under efficient diffusion of information:

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

- Multiple possible equilibria: depends on the initial beliefs.
- If everyone is uninformed and $p^H_{it} \approx 0 \; \forall it$, can targeting help?

Simulations: Methodology

- Generate random networks (characterized by G_{ij}s).
- Generate p^*_{iH}s to be correlated according to G_{ij}s (manipulate this to vary degree of heterogeneity).
- Select information entry points (initially $p^H_{it} \approx 0 \; \forall it$):
  - Centrality Based
  - Probability Based
- Let the diffusion take place for a few periods.
- Measure efficiency of a seeding strategy as the following:

$$\text{Targeting Efficiency} = \frac{\text{\% of informed households}}{\text{\% of informed households under full efficiency}}$$

- Repeat procedure for multiple networks.
- Evaluate results *on average*.

Simulation Results

Table 1: Simulation Results

		Level of Correlation			
Strategy	Statistic	Low	Medium	High	Perfect
Betweenness Centrality Based	Mean	0.72	0.82	0.84	0.75
	Variance	0.16	0.10	0.09	0.08
Probability Based	Mean	0.93	0.81	0.76	0.72
	Variance	0.11	0.11	0.14	0.09
	Observations	200	197	192	200

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

▸ Robustness w.r.t different centrality measure ▸ Robustness w.r.t higher population

Simulation Results

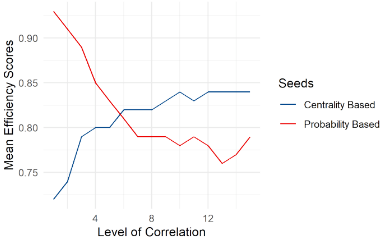


Figure 7: Mean efficiency scores over increasing levels of correlation

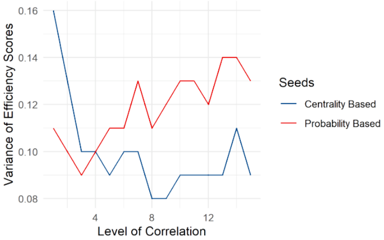


Figure 9: Variance of efficiency scores over increasing levels of correlation

Data

1. Replication data for Beaman et al., 2021 (BBMM):
   - RCT to promote pit planting (PP) for maize farmers in Malawi. Randomized information entry points at the village level.
   - Panel data contains information on adoption, demographics, and network characteristics. Timeline More Details

2. Agricultural Extension Services and Technology Adoption Survey (AESTAS) data collected by International Food Policy Research Institute (IFPRI).
   - Nationally representative survey of farmers in Malawi.
   - Panel data contains information on adoption of different technologies and household demographics. More Details

Identification

Using village level variations:

$$Y_{vt} = \beta_0 + \beta_1 Centrality_v + \beta_2 Probability_v + \beta_3 Het_v + \beta_4 Centrality_v \times Het_v + \beta_5 Probability_v \times Het_v + \lambda X_v + \zeta_t + \epsilon_{vt}$$

- Y_{vt}: adoption related outcome for village v at time t (excludes seed households).
- $Centrality_v$: average centrality of the seeds for village v at the baseline (available in the data).
- $Probability_v$: average probability of adoption of the seeds for village v at the baseline (not in the data). Approximation
- Het_v: coefficient of variation (CV) of probability of adoption at the village level.

Descriptive Statistics

Table 4a: Baseline Village-level Sample Characteristics

Variable	Mean (SD)
Adoption Rate (PP)	0.026 (0.060)
Any Non-Seed Adopters (PP)	0.345 (0.477)
Eigenvector Centrality of Seeds	0.182 (0.096)
Predicted Adoption Index of Seeds	0.101 (0.036)
Predicted Usage Index of Seeds	0.175 (0.035)
CV of Predicted Adoption Index	0.378 (0.071)
CV of Predicted Usage Index	0.187 (0.038)
Observations	200

Notes: Seed level measures are calculated using the average of two seeds, whenever the information on both seeds are available. Otherwise they reflect the information for one seed. Coefficient of Variations (CV) are calculated at the village level for the whole village. Adoption Rate and Any Non-Seed Adopters are calculated excluding seed or shadow farmers in a village.

Descriptive Statistics

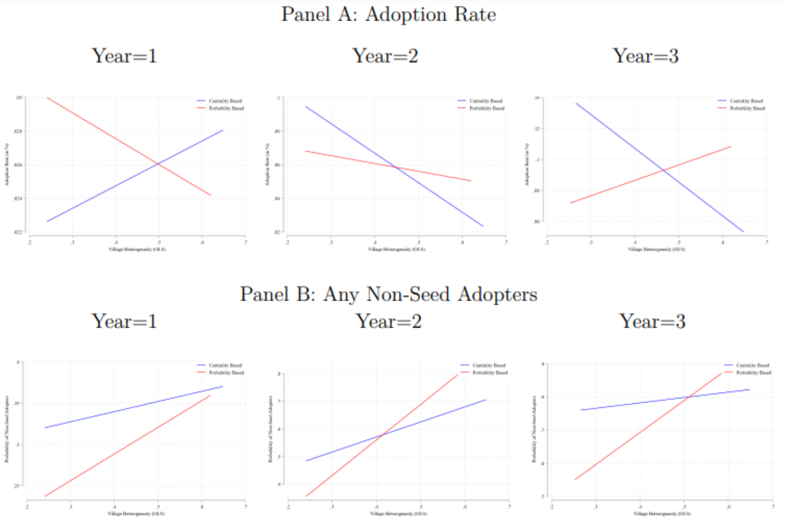


Figure 11: Outcomes for Different Seeding Strategies with respect to Village Heterogeneity

Regression Results

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

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

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

Discussion

Key Takeaway: Network-based targeting may require more than identifying central households within a social network.

⇒ We need to have an understanding of possible heterogeneity in benefits across households.

- Simulations show that centrality (probability) based targeting perform worse (better) as heterogeneity increase.
- Empirical results show support in favor of my hypotheses:
  - Positive (negative) effect of seeds' centrality (probability) on adoption decrease with increase in village-level heterogeneity. Robustness
  - Weaker evidences in favor of my hypotheses are found using the experimental variations in the data. Identification and Results

THANK YOU!

Theoretical Model

- Consider a two-stage decision process:
  - **Stage 1:** The households decide whether or not to make an irreversible investment to learn about an available new technology.
  - **Stage 2:** Conditional on making that investment, in the second stage they decide whether to stick to a traditional technology, or adopt the new technology.
- Traditional technology has a sure payoff of π^T, where the new technology provides a payoff of $\pi^N(\omega_{it})$, $\omega_{it} \in \Omega$.
- ω_{it} is drawn independently at each period t according to the true distribution $p_i^*(\omega_{it})$ for household i. Draws are not correlated over time within household and between households.

Theoretical Model

- But, true distributions are positively correlated between households according to the existing network structure (more details below).
- $\forall it$, $\exists \omega_{it}, \omega'_{it} \in \Omega$ such that $\pi^N(\omega_{it}) \geq \pi^T \geq \pi^N(\omega'_{it})$.
- $\mathcal{I}$ denotes the set of all households.
- $\exists i, j \in \mathcal{I}$ such that $\int_{\omega_{it} \in \Omega} p_i^*(\omega_{it}) \pi^N(\omega_{it}) - c_i \geq \pi^T$ and $\int_{\omega_{jt} \in \Omega} p_j^*(\omega_{jt}) \pi^N(\omega_{jt}) - c_j \leq \pi^T$, with c_i being the cost of new technology for household i.
- Initially all households are uninformed $\Rightarrow$ p_i^*s are unknown.
- The household i has beliefs $p_{it}(\omega_{it})$ over the distribution of ω_{it} at period t.

Theoretical Model

- At period t, uninformed household i has the option to become informed by putting effort $e_{it} \in \{0, 1\}$.
- If $e_{i\tau} = 1$, $e_{it} = 1 \; \forall t \geq \tau$.
- If $e_{it} = 1$, the household learns the true distribution $p_i^*(\omega_{it})$ at cost η_i. The cost of learning is incurred the first time the household gets informed only.
- If $e_{it} = 0$, no effort cost is incurred and the household uses DeGroot averaging to approximate the true distribution.
- Let G denote the $n \times n$ weighted, directed, and non-negative influence matrix ($n = |\mathcal{I}|$), where $G_{ij} \geq 0$ represents the weight i places on j's opinion (with $\sum_{j \in \mathcal{I}} G_{ij} = 1$).

Theoretical Model

- Then $\hat{p}_{it} = \sum_{j \in \mathcal{I}} G_{ij} p_{jt-1}$ denotes household i's approximation based on others' opinion following the DeGroot averaging.
- The true distributions are positively correlated between the households such that: $p_i^* = \sum_{j \in \mathcal{I}} G_{ij} p_j^*$.
- The belief of household i at period t:

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

- Assume that households need to be informed before they adopt: helps me explicitly capture the point when the households stop seeking information from their peers.
- Assume the households to be risk-neutral and myopic.

◂ Back

Simulation Robustness 1

Back

Table A1: Simulation Robustness (w.r.t different centrality measure)

		Level of Correlation			
Strategy	Statistic	Low	Medium	High	Perfect
Pagerank Centrality Based	Mean	0.86	0.96	0.99	0.98
	Variance	0.16	0.05	0.04	0.01
Probability Based	Mean	0.93	0.81	0.76	0.72
	Variance	0.11	0.11	0.14	0.09
	Observations	200	197	192	200

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

Simulation Robustness 2

◂ Back

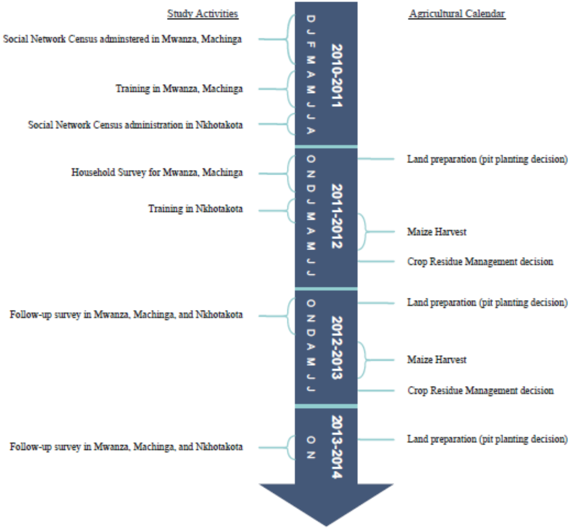
Table A2: Simulation Robustness (w.r.t different population)

		Level of Correlation			
Strategy	Statistic	Low	Medium	High	Perfect
Betweenness Centrality Based	Mean	0.59	0.55	0.54	0.55
	Variance	0.20	0.13	0.13	0.14
Probability Based	Mean	1.04	0.55	0.53	0.59
	Variance	0.06	0.14	0.14	0.13
	Observations	200	200	200	200

Notes: Simulations on varying levels of correlation are all done for 200 networks, each containing 50 households. However, upon generation of the true probabilities, some networks may be dropped as they contained 0% of informed households under full efficiency.

Timeline of BBMM

◂ Back



BBMM Replication Data

Back

- First collected the social network census data to elicit names of people each respondent consults when making agricultural decisions along with some other demographics.
- Used this responses with the village listing to identify links. Considered individuals linked if either party named each other or if they are part of the same household.
- Used simulations with the network information to identify seeds according to different diffusion processes to optimize diffusion after four periods.
- Randomly allocated villages to one of the four treatment groups and selected seeds for training based on that.
- Once the training is complete, randomly surveyed a panel of approximately 30 households per village, including all the seed and shadow farmers.

AESTAS Data

Back

- Objective was to monitor the *Lead Farmer* (LF) program in Malawi.
- Covers all districts of Malawi, except Likoma. Data collected in two waves: 2016 and 2018.
- Three types of interviews: Household, LF, and Community.
- Random sample of around 10 households were selected for interview from randomly selected sections within each district.
- Stratification was done based on whether or not the household had a LF.
- The same households were interviewed in the two waves with very small level of attrition (around 4%).
- For each household, both household head and their spouses were interviewed.

Approximating Probability of Adoption

◂ Back

- Proxy for probability of adoption using predicted adoption and usage indices. ▸ Construction of Indices
- Calculate these indices at the baseline, conditional on observable household demographics: number of adults and children, housing, livestock, and assets. ▸ Description of Variables
- Calculation uses estimates from following regressions using AESTAS data: $Adoption/Usage\ Index_{it} = f(X_{it}; \mu_{it})$. ▸ Results
- Based on a set of assumptions. ▸ All Assumptions

Calculating Adoption and Usage Indices

Back

- **Adoption Index:** Uses self-reported adoption for a list of pre-determined technologies and practices. This covered both agricultural and food processing practices. Average of these set of dummy variables taken to calculate the index.
- **Usage Index:** Self-reported plot-level usage for a list of pre-determined agricultural technologies and practices. Average of these set of dummy variables taken to calculate the index.

Description of Key Demographic Variables

Back

- **Adults**: Number of adults in the household.
- **Children**: Number of children in the household.
- **Housing**: Standardized first principal component (PC). Includes information on materials walls are made of, roof materials, floor materials (0- Traditional, 1- Modern), and whether the household has a toilet (only in the BBMM sample).
- **Livestock**: Standardized first PC. Includes the number of sheep, goats, chickens, cows, pigs the household owns. The BBMM sample also includes number of guinea fowl and doves.
- **Assets**: Standardized first PC. Includes the number of bicycles, radios and cell phones the household owns.

Approximating Probability: Assumptions

◂ Back

- **Assumption 1:** Adoption and Usage indices are good proxies for the probability of adoption.
- **Assumption 2:** The variation in adoption and usage indices, conditional on the observable demographics, is sufficient for my analysis. ▸ Actual and Predicted Variations
- **Assumption 3:** The mapping of observable characteristics to the adoption probability is the same across the datasets I use in this study. ▸ Sample Comparison
- **Assumption 4:** Any bias in the estimated relationship between adoption probability and observable characteristics is independent of the unobserved village-level learning in the BBMM sample.

Sample Comparison

Back

Table 2: Baseline Demographics Across Datasets

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

Notes: The variables *Adults* and *Children* represent number of adults and children in a household, respectively. The variables *Housing*, *Livestock*, and *Assets* were standardized first principal components. More details available in the paper.

Approximating Probabilities of Adoption

◂ Back

Table 3: OLS Regression Results for Adoption and Usage Indices

	Adoption Index		Usage Index	
Variables	(1)	(2)	(3)	(4)
Adults	0.008***	0.005**	0.011***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
Children	0.003***	0.002	0.003***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Housing	0.009***	0.007***	0.003	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Livestock	0.010***	0.005*	0.014***	0.009***
	(0.003)	(0.003)	(0.002)	(0.002)
Assets	0.024***	0.017***	0.020***	0.014***
	(0.002)	(0.002)	(0.002)	(0.002)
Household Controls	No	Yes	No	Yes
Observations	5610	5606	5610	5606
R-squared	0.096	0.149	0.085	0.123

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the section level are in parentheses. All regressions use a constant term and sample weights. The variables *Adults* and *Children* represent number of adults and children in a household, respectively. The variables *Housing*, *Livestock*, and *Assets* were standardized first principal components.

Approximating Probabilities of Adoption

Back

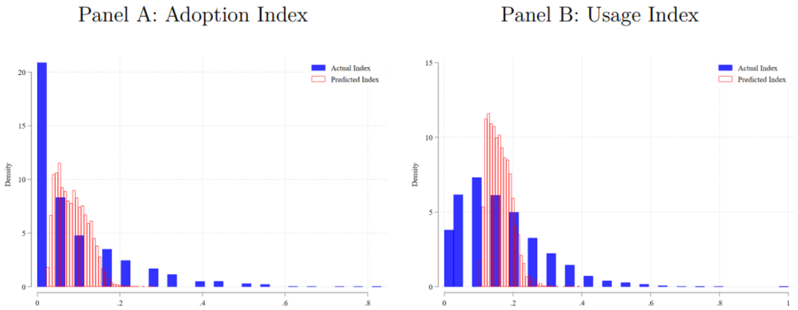


Figure 10: Actual and Predicted Adoption and Usage Indices

Regression Results

Table A1: Village level Regression 1 with Different Measure of Probability

Variables	Adoption Rate (1)	(2)	Any Non-Seed Adopters (3)	(4)
Eigenvector Centrality of Seeds	0.999*	0.817*	0.984	1.067
(=*Centrality*$_v$)	(0.565)	(0.480)	(1.302)	(1.191)
Predicted Usage Index of Seeds	-2.174	-1.511	-4.599	-0.0836
(=*Probability*$_v$)	(1.410)	(1.279)	(3.317)	(3.053)
CV of Predicted Usage Index	-1.091	-0.631	-2.549	2.142
(=*Heterogeneity*$_v$)	(0.805)	(0.779)	(2.905)	(2.823)
$Centrality_v \times Heterogeneity_v$	-4.481*	-3.936*	-4.874	-5.907
	(2.623)	(2.281)	(6.889)	(6.438)
$Probability_v \times Heterogeneity_v$	10.33*	7.276	23.13	0.889
	(6.160)	(5.623)	(14.19)	(13.40)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.063	0.174	0.037	0.164

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

Regression Results

◂ Back

Table A2: Village level Regression 1 with Different Measure of Centrality

Variables	Adoption Rate (1)	(2)	Any Non-Seed Adopters (3)	(4)
Closeness Centrality of Seeds	0.609**	0.454*	0.571	0.617
(=*Centrality*$_v$)	(0.306)	(0.234)	(0.709)	(0.659)
Predicted Adoption Index of Seeds	-2.438**	-1.709	-7.555**	-2.904
(=*Probability*$_v$)	(1.230)	(1.134)	(3.201)	(3.152)
CV of Predicted Adoption Index	-0.0774	-0.007	-0.677	0.887
(=*Heterogeneity*$_v$)	(0.214)	(0.202)	(1.196)	(1.158)
$Centrality_v \times Heterogeneity_v$	-1.325*	-1.020*	-1.552	-1.997
	(0.716)	(0.558)	(1.896)	(1.823)
$Probability_v \times Heterogeneity_v$	5.610**	3.814	17.55**	6.849
	(2.660)	(2.439)	(6.873)	(6.940)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.087	0.179	0.048	0.170

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

Identification Using Experimental Variation

Back

$$Y_{vt} = \psi_0 + \begin{bmatrix}\psi_1 & \psi_2\end{bmatrix} \begin{bmatrix} Centrality_v \\ Probability_v \end{bmatrix} + \psi_3 Het_v + \begin{bmatrix} 1 \\ Complex_v \\ Simple_v \\ Geo_v \end{bmatrix}' \Psi \begin{bmatrix} Centrality_v \\ Probability_v \end{bmatrix} Het_v + \gamma X_v + \rho_t + \eta_{vt}.$$

Effects are measured in terms of the Benchmark treatment:

- Same level of heterogeneity as the benchmark: Y_{vt} $\uparrow$es with centrality and $\downarrow$es with probability.
- Less heterogeneous: seeds with higher centrality perform better and seeds with higher probability perform worse. No prediction for seeds with less centrality and probability.
- Higher heterogeneity: seeds with lower centrality perform better and seeds with lower probability perform worse. No prediction for seeds with more centrality and probability.

Regression Results

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

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

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