Optimal Network-Based Targeting for Technology Adoption in Developing Countries

Aranya Chakraborty

McGill University

Table of Contents

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:** The diffusion only depends on the agents' positions in the network.

This Study

What if the agents differ in terms of other characteristics that affect the diffusion process?
$\Rightarrow$ If agents differ in terms of benefits from a new technology, how does that affect diffusion?

- Recommended network-based targeting strategies still optimal? If not, what is optimal in such a scenario?

This Study

What if the agents differ in terms of other characteristics that affect the diffusion process?
⇒ If agents differ in terms of benefits from a new technology, how does that affect diffusion?

- Recommended network-based targeting strategies still optimal? If not, what is optimal in such a scenario?

To answer these questions:

- Theoretically model agents 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.

Preview of Results

Simulations suggest that optimal targeting strategy depends on the population level heterogeneity.

- Targeting agents solely based on their positions in the network still works if the population is homogeneous in terms of their benefits.
- However, targeting early adopters works better if the population level heterogeneity is high.

Preview of Results

Simulations suggest that optimal targeting strategy depends on the population level heterogeneity.

- Targeting agents solely based on their positions in the network still works if the population is homogeneous in terms of their benefits.
- However, targeting early adopters works better if the population level heterogeneity is high.

Empirical analysis shows support in favor of the above findings.

- Central seeds lead to the less adoption as population heterogeneity increase.
- Seeds with higher probability of adoption lead to the less adoption as population heterogeneity increase.

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.

Table of Contents

Elements of the Model

- Two stage decision process: first learning, then adoption.
- Traditional technology has a sure payoff of π^T, where the new technology provides a payoff of $\pi^N(\omega_{it})$, $\omega_{it} \in \Omega$.
- Draws depend on the true distribution $p_i^*(\omega_{it})$ for household i. Independent draws every period.
- Uninformed households $\Rightarrow$ p_i^*s are unknown. Need to be fully informed (know p_i^*) before adoption.
- If uninformed, can become informed by putting effort $e_{it} \in \{0, 1\}$ at cost η_i.
- Costly effort: network ties help make this decision.
- Households are risk-neutral and myopic.

Elements of the model: Two-step adoption decision

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

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

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

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

▸ 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. Based on $\hat{p}_{it}$, 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 based on that, and $p_{is} = p_i^* \; \forall s \geq t$.

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?

Table of Contents

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. Illustrative Example

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. ▸ Illustrative Example
- 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

Strategy	Statistic	Level of Correlation: 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

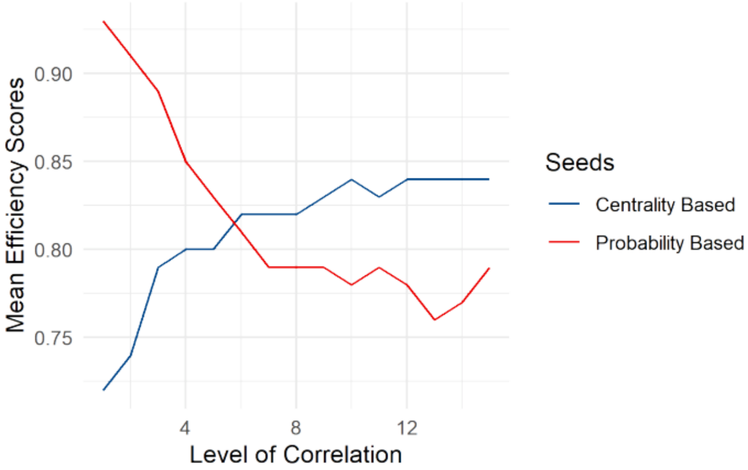


Figure 7: Mean efficiency scores over increasing levels of correlation

Simulation Results

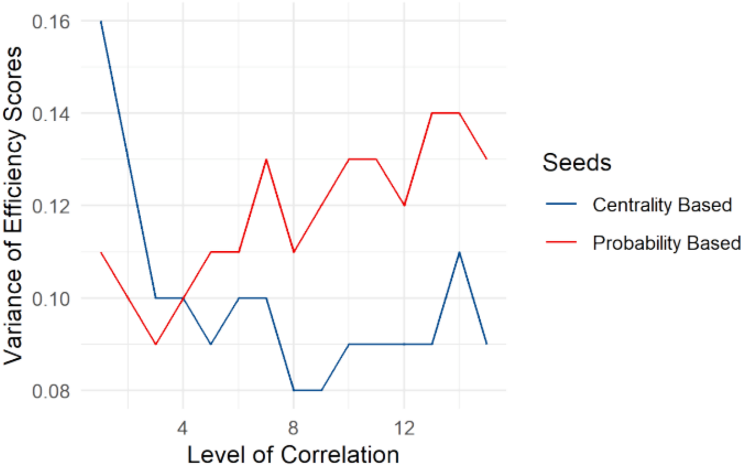


Figure 9: Variance of efficiency scores over increasing levels of correlation

Table of Contents

Hypotheses

Hypothesis 1: As the level of heterogeneity in terms of the benefits from a new technology ↑es, the success of central seeds in terms of diffusing that technology ↓es.

Hypotheses

Hypothesis 1: As the level of heterogeneity in terms of the benefits from a new technology ↑es, the success of central seeds in terms of diffusing that technology ↓es.

Hypothesis 2: As the level of heterogeneity in terms of the benefits from a new technology ↑es, the success of probability-based seeds in terms of diffusing that technology ↑es.

Data

1. Replication data for Beaman, BenYishay, Magruder, and Mobarak, 2021 (henceforth, 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 (henceforth, 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 for 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: Village-level Variations

Table 4: Baseline Village-level Sample Characteristics

	Treatment Status				
Variable	Benchmark	Complex	Simple	Geo	Overall
Adoption Rate (PP)	0.018	0.030	0.029	0.029	0.026
	(0.035)	(0.063)	(0.060)	(0.077)	(0.060)
Any Non-Seed Adopters (PP)	0.300	0.340	0.320	0.420	0.345
	(0.463)	(0.479)	(0.471)	(0.499)	(0.477)
Eigenvector Centrality of Seeds[†]	0.178	0.235	0.187	0.129	0.182
	(0.090)	(0.077)	(0.096)	(0.090)	(0.096)
Predicted Adoption Index of Seeds[‡]	0.110	0.114	0.101	0.082	0.101
	(0.034)	(0.036)	(0.041)	(0.025)	(0.036)
CV of Predicted Adoption Index	0.389	0.378	0.379	0.366	0.378
	(0.069)	(0.077)	(0.075)	(0.062)	(0.071)
Observations	50	50	50	50	200

Notes: † Contains 44 observations for the benchmark treatment group, 49 observations for the other treatment groups. 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: Village-level Variations

Table 4: Baseline Village-level Sample Characteristics

	Treatment Status				
Variable	Benchmark	Complex	Simple	Geo	Overall
Adoption Rate (PP)	0.018	0.030	0.029	0.029	0.026
	(0.035)	(0.063)	(0.060)	(0.077)	(0.060)
Any Non-Seed Adopters (PP)	0.300	0.340	0.320	0.420	0.345
	(0.463)	(0.479)	(0.471)	(0.499)	(0.477)
Eigenvector Centrality of Seeds†	0.178	0.235	0.187	0.129	0.182
	(0.090)	(0.077)	(0.096)	(0.090)	(0.096)
Predicted Adoption Index of Seeds‡	0.110	0.114	0.101	0.082	0.101
	(0.034)	(0.036)	(0.041)	(0.025)	(0.036)
CV of Predicted Adoption Index	0.389	0.378	0.379	0.366	0.378
	(0.069)	(0.077)	(0.075)	(0.062)	(0.071)
Observations	50	50	50	50	200

Notes: † Contains 44 observations for the benchmark treatment group, 49 observations for the other treatment groups. 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: Village-level Variations

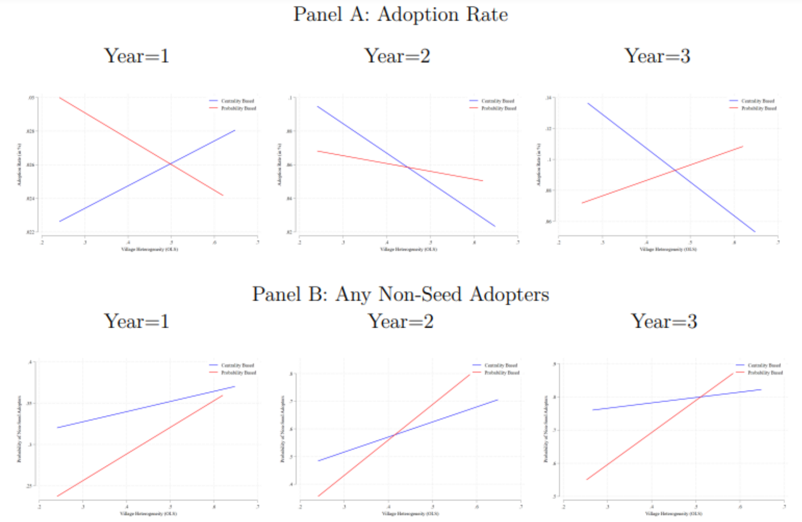


Figure 11: Outcomes for Different Seeding Strategies with respect to Village Heterogeneity

Regression Results: Village-level Variations

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.

Regression Results: Village-level Variations

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.

Regression Results: Village-level Variations

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.

Identification Using Experimental Variation

$$
\begin{aligned}
Y_{vt} = & \psi_0 + \psi_1 Cent_v + \psi_2 Prob_v + \psi_3 Het_v + \psi_4^0 Cent_v \times Het_v \\
& + \psi_4^T Cent_v \times Het_v \times Treat_v + \psi_5^0 Prob_v \times Het_v \\
& + \psi_5^T Prob_v \times Het_v \times Treat_v + \gamma X_v + \rho_t + \eta_{vt}.
\end{aligned}
$$

- $Treatment_v$: captures whether the village v belongs to complex, simple or geo treatment arm.
- Effects are measured in terms of the omitted category (benchmark treatment arm).
- Villages are less (or, same level of) heterogeneous in other treatment arms (compared to benchmark). That implies:
  - Y_{vt} $\uparrow$es with centrality and $\downarrow$es with probability.
  - No prediction for seeds with less centrality and probability.

Descriptive Statistics: Experimental Variations

Table 4: Baseline Village-level Sample Characteristics

	Treatment Status				
Variable	Benchmark	Complex	Simple	Geo	Overall
Adoption Rate (PP)	0.018 (0.035)	0.030 (0.063)	0.029 (0.060)	0.029 (0.077)	0.026 (0.060)
Any Non-Seed Adopters (PP)	0.300 (0.463)	0.340 (0.479)	0.320 (0.471)	0.420 (0.499)	0.345 (0.477)
Eigenvector Centrality of Seeds†	0.178 (0.090)	0.235 (0.077)	0.187 (0.096)	0.129 (0.090)	0.182 (0.096)
Predicted Adoption Index of Seeds‡	0.110 (0.034)	0.114 (0.036)	0.101 (0.041)	0.082 (0.025)	0.101 (0.036)
CV of Predicted Adoption Index	0.389 (0.069)	0.378 (0.077)	0.379 (0.075)	0.366 (0.062)	0.378 (0.071)
Observations	50	50	50	50	200

Notes: † Contains 44 observations for the benchmark treatment group, 49 observations for the other treatment groups. 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: Experimental Variations

Table 4: Baseline Village-level Sample Characteristics

	Treatment Status				
Variable	Benchmark	Complex	Simple	Geo	Overall
Adoption Rate (PP)	0.018	0.030	0.029	0.029	0.026
	(0.035)	(0.063)	(0.060)	(0.077)	(0.060)
Any Non-Seed Adopters (PP)	0.300	0.340	0.320	0.420	0.345
	(0.463)	(0.479)	(0.471)	(0.499)	(0.477)
Eigenvector Centrality of Seeds†	0.178	0.235	0.187	0.129	0.182
	(0.090)	(0.077)	(0.096)	(0.090)	(0.096)
Predicted Adoption Index of Seeds‡	0.110	0.114	0.101	0.082	0.101
	(0.034)	(0.036)	(0.041)	(0.025)	(0.036)
CV of Predicted Adoption Index	0.389	0.378	0.379	0.366	0.378
	(0.069)	(0.077)	(0.075)	(0.062)	(0.071)
Observations	50	50	50	50	200

Notes: † Contains 44 observations for the benchmark treatment group, 49 observations for the other treatment groups. 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: Experimental Variations

Table 4: Baseline Village-level Sample Characteristics

	Treatment Status				
Variable	Benchmark	Complex	Simple	Geo	Overall
Adoption Rate (PP)	0.018	0.030	0.029	0.029	0.026
	(0.035)	(0.063)	(0.060)	(0.077)	(0.060)
Any Non-Seed Adopters (PP)	0.300	0.340	0.320	0.420	0.345
	(0.463)	(0.479)	(0.471)	(0.499)	(0.477)
Eigenvector Centrality of Seeds†	0.178	0.235	0.187	0.129	0.182
	(0.090)	(0.077)	(0.096)	(0.090)	(0.096)
Predicted Adoption Index of Seeds‡	0.110	0.114	0.101	0.082	0.101
	(0.034)	(0.036)	(0.041)	(0.025)	(0.036)
CV of Predicted Adoption Index	0.389	0.378	0.379	0.366	0.378
	(0.069)	(0.077)	(0.075)	(0.062)	(0.071)
Observations	50	50	50	50	200

Notes: † Contains 44 observations for the benchmark treatment group, 49 observations for the other treatment groups. 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.

Regression Results: Experimental Variations

Table 6: Village level Regression 2 of Adoption Outcomes

Variables	Adoption Rate (5)	Adoption Rate (6)	Any Non-Seed Adopters (7)	Any Non-Seed Adopters (8)
$Centrality_v \times Heterogeneity_v$	-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	-0.070
	(0.940)	(0.839)	(2.235)	(2.098)
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 seed centrality, seed probability, village-level heterogeneity, 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: Experimental Variations

Table 6: Village level Regression 2 of Adoption Outcomes

Variables	Adoption Rate (5)	(6)	Any Non-Seed Adopters (7)	(8)
$Centrality_v \times Heterogeneity_v$	-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	-0.070
	(0.940)	(0.839)	(2.235)	(2.098)
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 seed centrality, seed probability, village-level heterogeneity, 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: Experimental Variations

Table 6: Village level Regression 2 of Adoption Outcomes

Variables	Adoption Rate (5)	(6)	Any Non-Seed Adopters (7)	(8)
$Centrality_v \times Heterogeneity_v$	-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	-0.070
	(0.940)	(0.839)	(2.235)	(2.098)
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 seed centrality, seed probability, village-level heterogeneity, 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: Experimental Variations

Table 6: Village level Regression 2 of Adoption Outcomes

Variables	Adoption Rate (5)	(6)	Any Non-Seed Adopters (7)	(8)
Centrality$_v$ × *Heterogeneity*$_v$	-2.423**	-2.237**	-6.692	-6.574
	(1.093)	(0.996)	(4.503)	(4.119)
Centrality$_v$ × *Heterogeneity*$_v$ × *Complex*	0.657**	0.664**	4.328**	3.756**
	(0.306)	(0.282)	(1.775)	(1.664)
Centrality$_v$ × *Heterogeneity*$_v$ × *Simple*	0.416	0.428	1.078	0.431
	(0.337)	(0.320)	(2.060)	(1.947)
Centrality$_v$ × *Heterogeneity*$_v$ × *Geo*	2.026**	1.942**	0.103	-0.070
	(0.940)	(0.839)	(2.235)	(2.098)
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 seed centrality, seed probability, village-level heterogeneity, 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

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

Variables	Adoption Rate (5)	Adoption Rate (6)	Any Non-Seed Adopters (7)	Any Non-Seed Adopters (8)
$Probability_v \times Heterogeneity_v$	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 seed centrality, seed probability, village-level heterogeneity, 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

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

Variables	Adoption Rate (5)	(6)	Any Non-Seed Adopters (7)	(8)
$Probability_v \times Heterogeneity_v$	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 seed centrality, seed probability, village-level heterogeneity, 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

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

Variables	Adoption Rate (5)	(6)	Any Non-Seed Adopters (7)	(8)
$Probability_v \times Heterogeneity_v$	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 seed centrality, seed probability, village-level heterogeneity, 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

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

Variables	Adoption Rate (5)	(6)	Any Non-Seed Adopters (7)	(8)
Probability$_v$ × *Heterogeneity*$_v$	5.881**	4.104*	22.97***	12.35
	(2.437)	(2.286)	(7.720)	(7.626)
Probability$_v$ × *Heterogeneity*$_v$ × *Complex*	-0.155	-0.232	-1.275	-0.679
	(0.520)	(0.497)	(2.765)	(2.654)
Probability$_v$ × *Heterogeneity*$_v$ × *Simple*	-0.121	-0.110	1.941	3.511
	(0.642)	(0.571)	(3.572)	(3.333)
Probability$_v$ × *Heterogeneity*$_v$ × *Geo*	-2.588**	-2.562**	-0.391	0.538
	(1.131)	(1.039)	(4.028)	(3.618)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.133	0.224	0.113	0.222

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

Table of Contents

Summary

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. ▸ Robustness

Remarks

Difficulty in targeting based on the probability of adoption.

- Use of additional data to predict adoption conditional on observable demographics.
- Better approach: collect additional information while collecting network data.

Remarks

Difficulty in targeting based on the probability of adoption.

- Use of additional data to predict adoption conditional on observable demographics.
- Better approach: collect additional information while collecting network data.

Policy: Need to have an understanding of possible heterogeneity in benefits across households.

- Demands more information.
- Targeting vs. random seeding: cost-benefit analysis needed.

THANK YOU!

Theoretical Model

- Two-stage decision process:
  - **Stage 1:** Households decide whether or not to make an irreversible investment to learn about an available new technology.
  - **Stage 2:** Conditional on the investment, 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

Illustrative example

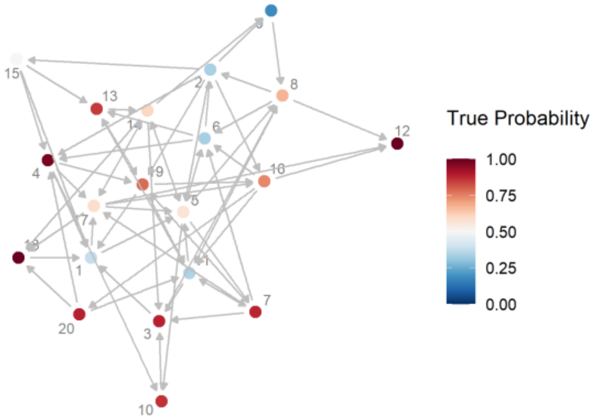


Figure 1: Distribution of True Probability within the network

Illustrative example

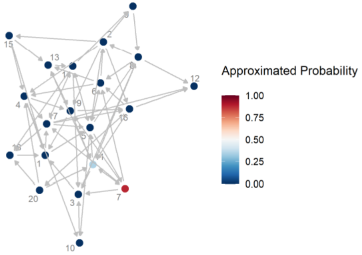


Figure 2: Seeding based on Centrality

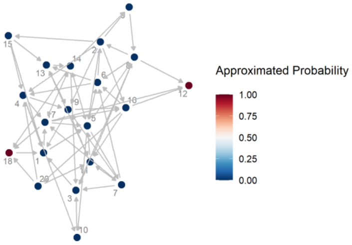


Figure 3: Seeding based on Probability

Illustrative example

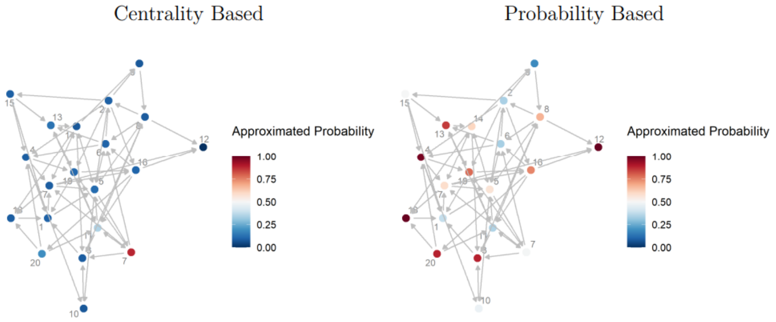


Figure 4: Performance of seeds after 10 periods

Back

Simulation Robustness 1

◂ Back

Table D.7: 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

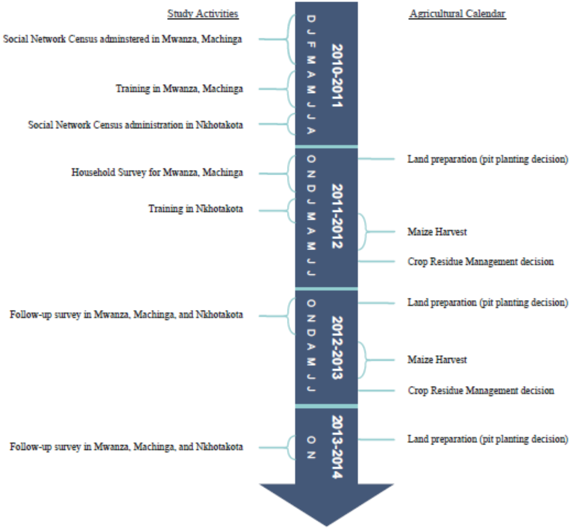
Table D.8: Simulation Robustness (w.r.t different population)

		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

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

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 probability of adoption.
- **Assumption 2:** The variation in adoption and usage indices that can be captured by the observable demographics, are sufficient for my analysis. ▸ Actual and Predicted Variations
- **Assumption 3:** The mapping from the observable demographics to the adoption and usage indices are the same in the BBMM sample, as it is in the AESTAS sample. ▸ Sample Comparison

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
	Skewness	2.59	1.00	0.54	3.74	0.31
	Kurtosis	16.76	6.01	2.01	26.67	1.79
	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
	Skewness	1.21	0.76	2.48	4.64	1.24
	Kurtosis	5.43	4.43	8.70	35.03	5.64
	Observations	5384	5407	5382	5407	5407

Approximating Probabilities of Adoption

◂ Back

Table 3: OLS Regression Results for Adoption and Usage Indices

Variables	Adoption Index (1)	(2)	Usage Index (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

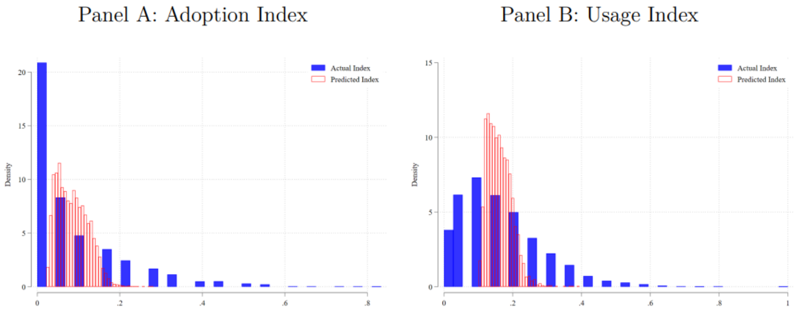


Figure 10: Actual and Predicted Adoption and Usage Indices

Regression Results

Table D.13: Village level Regression 1 with Different Measure of Probability

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

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

Regression Results

◂ Back

Table D.15: 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	17.550**	6.849
	(2.660)	(2.439)	(6.873)	(6.940)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.087	0.179	0.048	0.170

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

Regression Results

Table D.14: Village level Regression 2 with Different Measure of Probability

Variables	Adoption Rate (1)	(2)	Any Non-Seed Adopters (3)	(4)
$Centrality_v \times Heterogeneity_v$	-4.619*	-4.617*	-12.420	-12.190
	(2.549)	(2.473)	(8.555)	(7.660)
$Centrality_v \times Heterogeneity_v \times Complex$	1.432*	1.595**	9.431**	8.099**
	(0.749)	(0.720)	(4.323)	(3.996)
$Centrality_v \times Heterogeneity_v \times Simple$	0.492	0.576	3.308	1.958
	(0.860)	(0.831)	(4.665)	(4.340)
$Centrality_v \times Heterogeneity_v \times Geo$	3.957*	3.711**	-1.692	-2.661
	(2.057)	(1.785)	(4.676)	(4.495)
$Probability_v \times Heterogeneity_v$	10.260*	7.702	33.700*	13.410
	(5.561)	(5.378)	(17.390)	(16.260)
$Probability_v \times Heterogeneity_v \times Complex$	-0.316	-0.589	-2.606	-1.839
	(0.762)	(0.778)	(4.577)	(4.315)
$Probability_v \times Heterogeneity_v \times Simple$	0.428	0.416	1.355	3.119
	(0.984)	(0.866)	(5.269)	(4.868)
$Probability_v \times Heterogeneity_v \times Geo$	-2.468*	-2.409**	2.565	3.786
	(1.377)	(1.217)	(4.925)	(4.505)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.114	0.212	0.100	0.215

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors are in parentheses. All regressions include seed centrality, seed probability, village-level heterogeneity, 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 D.16: Village level Regression 2 with Different Measure of Centrality

Variables	Adoption Rate (1)	(2)	Any Non-Seed Adopters (3)	(4)
$Centrality_v \times Heterogeneity_v$	-1.457**	-1.181**	-2.508	-3.114
	(0.591)	(0.478)	(1.935)	(1.939)
$Centrality_v \times Heterogeneity_v \times Complex$	0.307**	0.304**	1.446*	1.355*
	(0.137)	(0.140)	(0.838)	(0.810)
$Centrality_v \times Heterogeneity_v \times Simple$	0.364**	0.395***	-0.401	-0.498
	(0.157)	(0.152)	(0.934)	(0.917)
$Centrality_v \times Heterogeneity_v \times Geo$	0.679**	0.667**	0.517	0.140
	(0.267)	(0.262)	(0.988)	(0.914)
$Probability_v \times Heterogeneity_v$	4.791**	3.306	19.310***	9.942
	(2.281)	(2.166)	(7.105)	(6.963)
$Probability_v \times Heterogeneity_v \times Complex$	-0.351	-0.419	0.056	0.189
	(0.632)	(0.637)	(3.155)	(3.031)
$Probability_v \times Heterogeneity_v \times Simple$	-1.125*	-1.235*	4.299	5.406
	(0.664)	(0.629)	(3.876)	(3.727)
$Probability_v \times Heterogeneity_v \times Geo$	-2.855**	-2.864**	-2.748	0.060
	(1.200)	(1.187)	(4.867)	(4.398)
Village-level Controls	No	Yes	No	Yes
Observations	324	324	324	324
R-squared	0.121	0.209	0.109	0.223

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors are in parentheses. All regressions include seed centrality, seed probability, village-level heterogeneity, 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.