

Can Network Theory-Based Targeting Increase Technology Adoption?[†]

By LORI BEAMAN, ARIEL BENYISHAY, JEREMY MAGRUDER,
AND AHMED MUSHFIQ MOBARAK*

Can targeting information to network-central farmers induce more adoption of a new agricultural technology? By combining social network data and a field experiment in 200 villages in Malawi, we find that targeting central farmers is important to spur the diffusion process. We also provide evidence of one explanation for why centrality matters: a diffusion process governed by complex contagion. Our results are consistent with a model in which many farmers need to learn from multiple people before they adopt themselves. This means that without proper targeting of information, the diffusion process can stall and technology adoption remains perpetually low. (JEL O13, O18, O33, Q12, Q16)

Technology diffusion is critical for growth and development (Alvarez, Buera, and Lucas 2013; Perla and Tonetti 2014). Information frictions are potential constraints to technology adoption, and social relationships can serve as important vectors through which individuals learn about, and are then convinced to adopt, new technologies.¹ Adoption of apparently productive new technologies has often been frustratingly slow (Ryan and Gross 1943; Munshi 2007; Jack 2011; Qiao, Huang, and Wang 2015). This generates a policy priority: how can policymakers effectively use social relationships to promote technological diffusion? In this paper, we implement a field experiment in which we choose entry points of information (seeds) into

*Beaman: Department of Economics and Institute for Policy Research, Northwestern University (email: l-beaman@northwestern.edu); BenYishay: Department of Economics, College of William and Mary (email: abenyishay@wm.edu); Magruder: Agricultural and Resource Economics, University of California-Berkeley (email: jmagruder@berkeley.edu); Mobarak: Yale University (ahmed.mobarak@yale.edu). Esther Duflo was the coeditor for this article. We thank the CEGA/JPAL Agricultural Technology Adoption Initiative (ATAI) and 3ie for financial support. Beaman acknowledges support by the National Science Foundation under grant 1254380. We gratefully acknowledge the support and cooperation of Paul Fatch, Readwell Musopole, and many other staff members of the Malawi Ministry of Agriculture. Thomas Coen, Niall Kelleher, Maria Jones, Ofer Cohen, Allen Baumgardner-Zuzik, and the IPA-Malawi country office provided invaluable support for data collection. Hossein Alidaee provided excellent research assistance. We thank, without implicating, the anonymous referees, Arun Chandrasekhar, Matt Jackson, Kaivan Munshi, Chris Udry and numerous seminar audiences for very helpful comments. The study is registered with the AEA registry under AEARCTR-0002017 (see Beaman et al. 2018). IRB approval was provided by Northwestern University (STU00030251) and MIT (COUHES 1005003884). The data repository is openicpsr-130605 (see Beaman et al. 2021). All errors are our own.

[†]Go to <https://doi.org/10.1257/aer.20200295> to visit the article page for additional materials and author disclosure statements.

¹Large literatures in economics (Duflo and Saez 2003, Munshi 2008, Magruder 2010, Beaman 2012), finance (Bursztyn et al. 2014), sociology (Rogers 1962), and medicine and public health (Coleman, Katz, and Menzel 1957; Flodgren et al. 2007; Oster and Thornton 2012) show that information and behaviors spread through interpersonal ties.

a social network and introduce a productive new agricultural technology via those seeds across 200 villages in Malawi.

A rich empirical literature has documented faster diffusion when technologies were seeded with people who are central in the network (Banerjee et al. 2013 in the context of microfinance in India; Banerjee et al. 2019 in the context of immunization in India; Kim et al. 2015 looking at health behaviors in Honduras). Targeting information to central agents in a network can even work better than broadcasting information widely (Banerjee et al. 2020).

These empirical patterns that establish the importance of centrality may be surprising given recent theoretical discussion by Akbarpour, Malladi, and Saberi (2020)—henceforth, AMS—which shows that in many canonical diffusion models, adding a few additional seeds leads to more diffusion than targeting central people to serve as seeds. The class of models AMS consider require three conditions. First, agents must adopt a new behavior after a single exposure to someone else who has adopted in the network. This is called “simple contagion” and is the base for work-horse models like the Susceptible-Infected-Recovered (SIR) model. Second, the time period for adoption is sufficiently long. And finally, social interaction within the network is frequent. The intuition for the AMS result is straightforward: whether central or not, people are connected to their local network, so given enough time and enough talk, messages will spread through their connections and quickly reach the well-connected people at the network’s center. Adding a few more seeds at random increases the probability that at least one of them will be close to the well-connected center to begin with, making targeting relatively unimportant. However, if any of the three criteria fail, then targeting may be necessary to prevent information frictions from curbing widespread technological diffusion.

Our paper helps bridge the gap between these theoretical and empirical results. We implemented a randomized controlled trial where we used different variants of the threshold model of diffusion (e.g., Granovetter 1978; Centola and Macy 2007; Acemoglu, Ozdaglar, and Yildiz 2011) to choose seeds. This creates a unifying framework which both generates variation in seed centrality across treatment arms and also helps us explore why targeting may matter for technology diffusion.

Our experiment takes place in an important real-world context: agricultural extension services in developing countries. Agricultural productivity growth in Africa has stalled (World Bank 2008), in part because of a slow adoption rate of new technologies. Agricultural extension is the key policy tool governments use to promote technology adoption (Anderson and Feder 2007), and it often relies on social learning.² We partnered with the Ministry of Agriculture in Malawi to run an experiment that could enhance the effectiveness of its extension services by partnering with two “seed” farmers in each study village who could induce widespread social learning. The experiment was implemented in 200 villages, with 50 villages in each of the 4 treatment groups. The specific technology promoted, “pit planting,” has the potential to significantly improve maize yields in arid areas of rural Africa.³

²A large literature has established that social learning about agricultural practices influences the uptake of new technologies among farmers (Griliches 1957, Foster and Rosenzweig 1995, Munshi 2004, Bandiera and Rasul 2006, Conley and Udry 2010, Burlig and Stephens 2019, Islam et al. 2019).

³It has been shown to increase productivity by 40–100 percent in tests conducted under controlled conditions (Haggblade and Tembo 2003); in large-sample field tests conducted under realistic “as implemented by

It is a practice that was largely unknown in Malawi, and learning is therefore crucial for the diffusion of this technology.

In the Benchmark treatment, extension agents chose the seeds as they normally would (status quo or picking by experts). In the remaining treatment groups, we strategically chose the seeds using detailed social network data we collected in every village. We ensured that selected seeds in different treatments would inhabit different parts of the network by exploiting variations on the threshold model of diffusion to suggest pairings of seeds that may be more or less effective, given different underlying diffusion processes. In the second treatment group, we selected seeds who would (in theoretical simulations) optimize diffusion over a 4-year period, if the diffusion process is characterized by a complex contagion. Complex contagion is a diffusion process in which technology only diffuses when individuals are connected to at least two knowledgeable farmers. The pair of seeds chosen by this complex contagion treatment are both central in the network. Seed selection in the third treatment is the result of simulations of the simple contagion variant of the threshold model, where farmers only need to know one knowledgeable farmer. In simple diffusion, a single central seed will diffuse to the dense part of the network so that a second seed is best used to diffuse to the more distant periphery. As a result, one of the seeds is network-central while the second person is typically more peripheral. This variant of the model is similar to those considered in the AMS framework. In the final treatment group, we used geography to proxy for social network data, to create a cheaper, “scalable” approach coupled with the complex contagion model. These seeds are typically low centrality, but are close to each other in the network.

During the 3-year period of the experiment, pit planting adoption grew from 0 percent to about 11 percent in the villages with two highly central seeds. This rate of increase in adoption is comparable to the spread of some very profitable new agricultural technologies (e.g., Munshi 2007). Ryan and Gross (1943) show that it took 10 years for hybrid seed corn to be adopted in Iowa in the 1930s. The adoption rate is 3 percentage points lower in Benchmark villages in years 2 and 3, though only the year 2 differences are statistically significant.

We also test whether the initial advantage of central seeding will likely dissipate over time by examining another important metric: whether any farmers in the village other than the seeds adopt. If there is no diffusion within the first three years, it is unlikely that conversation and experience over longer time horizons will inspire broad technology adoption. We observe a critical failure of expert-based seeding. There is no diffusion of pit planting in 45 percent of the Benchmark villages after 3 years. In villages with two highly central seeds, there was a 56 percent greater likelihood ($p < 0.01$) that at least one person other than the seeds adopts the technology in the village, relative to the Benchmark. The results clearly indicate that targeting central seeds was necessary to generate adoption of pit planting in Malawi.

We then turn to understanding *why* central targeting was so important in this context. One potential explanation is that the variant of the threshold model that we used to select seeds captures the underlying diffusion process. AMS and Jackson and

government” conditions (BenYishay and Mobarak 2019); and using experimental variation among villagers in the present study.

Storms (2019) demonstrate that targeting on the basis of centrality is more important when there is complex contagion. We show that different thresholds for technology adoption are naturally micro-founded through a naïve Bayesian learning model, as we discuss in Section IVA. We anticipate that learning about a new agricultural technology in a developing country is precisely a context in which agents may have a high threshold. This fact would have clear policy relevance: if farmers need to learn from more than one informed connection before they themselves adopt, this would generate a very slow and in many cases permanently stalled adoption pattern, just as we observe in Benchmark villages. Overcoming this problem would necessitate targeting central individuals, as in Banerjee et al (2019).

The diffusion we observe demonstrates several empirical regularities consistent with complex contagion. Though, we note that it is difficult to differentiate complex contagion from other reasons that targeting multiple central farmers may improve technology adoption. We observe three patterns suggested by complex contagion. First, a key insight from the threshold model is that poor targeting could lead to a complete failure of adoption within the village, as we see in our data. Second, consistent with our theory, we show that treatment effects are largest (i) in villages where there is more to learn, because baseline knowledge was lowest, and (ii) among farmers whose land is most suited to pit planting. Third, we use our farmer-level data to provide direct evidence in support of complex contagion. We leverage the random treatment assignment to identify that farmers who are connected to two seeds are more likely to learn about and adopt pit planting than farmers connected to only one seed, holding network position constant.

The targeting method used in this paper is a proof of concept, relying on an expensive method of collecting network data. As such, it is not intended to be practical or directly scalable. The next step is to use cheaper ways to identify highly central individuals. One could use gossips (Banerjee et al. 2019), cell phone data (Björkegren 2018; Blumenstock, Chi, and Tan 2019), or other administrative data (Bennett and Bergman 2020), or aggregated relational data from a sample of individuals (Breza et al. 2020) to achieve this.⁴

The rest of the paper is organized as follows. We start with the experimental setting and design, along with details on the implementation of the intervention. Section II describes the data. Section III presents the average treatment effects on pit planting adoption. In Section IV, we propose a theoretical model to explain the results, and provide supplemental evidence of the proposed mechanism. Section V discusses cost-effective and policy-relevant alternatives to the data-intensive network theory-based procedures we used in this paper, and discuss other options available in the literature. Section VI concludes.

⁴In our paper we did one lower cost method, the geography-based targeting strategy. It generated some gains in adoption relative to the Benchmark. However, physical proximity does not appear to be a good proxy for social connections in this context. A variety of other papers test the ability of local institutions, such as nominations or focus groups, to identify useful partners: Kremer et al. (2011) identify and recruit “ambassadors” to promote water chlorination in rural Kenya; Miller and Mobarak (2015) first market improved cookstoves to “opinion leaders” in Bangladeshi villages before marketing to others; and BenYishay and Mobarak (2019) incentivize “lead farmers” and “peer farmers” to partner with agricultural extension officers in Malawi. We also develop an intuitive algorithm to identify central farmers that can be implemented with a small number of interviews, and simulations on our data show that this method would generate large gains in technology adoption.

I. Field Experiment

A. Setting

Our experiment on technology diffusion within an agricultural extension system takes place in 200 villages randomly sampled from 3 Malawian districts with largely semi-arid climates (Machinga, Mwanza, and Nkhotakota). Approximately 80 percent of Malawi's population lives in rural areas (World Bank 2011), and agricultural production in these areas is dominated by maize: 97 percent of farmers grow maize, and over one-half of households grow no other crop (Lea and Hanmer 2009). Technology adoption and productivity in maize is thus closely tied to welfare.

The existing agricultural extension system in Malawi relies on Agricultural Extension Development Officers, henceforth extension agents, who are employed by the Ministry of Agriculture and Food Security (MoAFS). Many extension agents are responsible for upward of 30–50 villages, which implies that direct contact with villagers is rare. According to the 2006/2007 Malawi National Agricultural and Livestock Census, only 18 percent of farmers participate in any type of extension activity. Extension agents cope with these staff shortages by relying on a small number of lead farmers, who are trained, but not incentivized, to disseminate knowledge via social learning. Against this backdrop of staff shortages, maximizing the reach of social learning in the diffusion process may be a cost-effective way to improve the effectiveness of extension.

B. Experimental Design

We partner with the Malawi Ministry of Agriculture to select the appropriate technologies to promote and engage extension staff to train exactly two seed farmers in each study village. Our experimental variation only changes how those seed farmers are chosen and holds all other aspects of the training constant.

The experiment has four treatment arms. The Benchmark treatment is the status quo Benchmark, where extension agents were asked to select two seed farmers as they normally would in settings outside the experiment. In the remaining three treatment groups, we strategically chose the seeds to ensure that partner farmers were located in different parts of the network.

We identified farmers with different centrality characteristics in each of the study villages by choosing partners who would be the “theoretically optimal” choices as seeds under alternative formulations of the threshold model (e.g., Granovetter 1978; Centola and Macy 2007; Acemoglu, Ozdaglar, and Yildiz 2011). The threshold model of diffusion postulates that individuals adopt a behavior only if they are connected to at least a threshold number of adopters (λ).⁵

⁵In Section IVA, we will present a micro-foundation which demonstrates how a learning model can generate thresholds. In this version of the model, the threshold is based on the number of people informed about the technology, as opposed to the number of adopters directly.

The three treatment arms in which we selected the seeds using the threshold model are as follows:⁶

- (i) **Complex Contagion:** This treatment identified seeds by maximizing simulated diffusion when $E[\lambda] \approx 2$ using network relationship data. The two selected seeds are usually both very central in the network.
- (ii) **Simple Contagion:** This treatment identified seeds by maximizing simulated diffusion when $E[\lambda] \approx 1$ using network relationship data. In most networks, this identifies one seed who is central and one who is not.
- (iii) **Geo Treatment:** This treatment typically identifies two seeds who are near each other in the network, but are not be central. This resulted from maximizing simulated diffusion when $E[\lambda] \approx 2$ using network data constructed using only geographic proximity.

The intuition for why the different formulations of the threshold model generates these different targeting strategies is as follows. When many farmers have a threshold for adoption above 1, what this literature calls *complex contagion*, targeting becomes essential because one needs to seed information in part of the network that is dense and where the seeds have connections in common. In this model, identifying two seeds who are both central to the network is important for diffusion.⁷ In contrast, when the threshold is generally equal to 1, what the literature calls *simple contagion*, identifying a single seed in the central part of the network is sufficient to achieve widespread diffusion. In this case, a second seed is optimally located in a more distant part of the network, so that both the center and the periphery can achieve quick take-up. Identifying the optimal seeds in each of these cases requires rich network data, described in Section II. We also implemented a fourth treatment, “Geo,” which substitutes household locations for the network graph under the assumption that nearby households are likely to be connected.

In online Appendix Section A.1, we discuss in detail the algorithm used to choose the seeds. Note that in all villages, we can construct which farmers would have been chosen as Simple diffusion seeds, Complex diffusion seeds, or Geo seeds, irrespective of the village’s assigned treatment condition. We call the counterfactual seed farmers “shadow” farmers. We also use the term “partner” to refer to an individual who would be a Simple, Complex, or Geo seed irrespective of whether they are

⁶In other words, we randomly assign the “threshold model formulations” to different villages. Randomization was stratified by district, and implemented using a re-randomization procedure which checked balance on the following covariates: percent of village using compost at baseline; percent village using fertilizer at baseline; and percent of village using pit planting at baseline. Randomization was implemented in each district separately.

⁷As we will see later, this feature has significant ramifications for targeting: while randomly selected seeds are quite likely to be relatively close to the center of any network, groups of randomly selected seeds remain unlikely to share ties in common.

TABLE 1—CENTRALITY OF PARTNER FARMERS ACROSS TREATMENTS

	Eigenvector centrality		Degree	
	Rank 1 partner (1)	Rank 2 partner (2)	Rank 1 partner (3)	Rank 2 partner (4)
<i>Treatment arms</i>				
Complex diffusion	0.28	0.19	17.49	13.39
Simple diffusion	0.27	0.07	16.59	6.70
Geographic	0.15	0.10	9.48	6.34
Benchmark	0.21	0.13	13.29	9.80

Notes: The sample includes all partner farmers, including seeds and shadows. However, Benchmark partners are restricted to only seed farmers (and hence the sample size is smaller) because Benchmark shadow farmers are not observed in Complex, Simple, or Geo villages.

trained and therefore become a seed.⁸ We do not observe shadow farmers for the Benchmark treatment.⁹

Table 1 demonstrates the centrality of the two selected partner farmers in each treatment arm. The most central of the two partners (Rank 1 partners), as measured by eigenvector centrality¹⁰ in column 1, is similarly central in both the Complex and Simple diffusion, but less central in Geo. However, the second partner highlights the key difference in the treatments. The second partner in the Complex treatment is much more central than the second partner in the Simple diffusion treatment. In Geo, neither partner is very central, but they are similarly central, highlighting that geography in this context was not a good proxy for social connectedness but that the targeting strategy was *ex ante* similar to the Complex diffusion strategy. If we use an alternative measure of social connectedness, degree (the number of contacts a person has) we see a similar pattern. Both Complex partners have many connections. The most connected partner in the Simple diffusion treatment is similar in the number of contacts to the most connected partner in Complex diffusion, but the second partner is much less connected.

The Benchmark seeds, which were chosen by extension officers using their own criteria, show an intermediate level of centrality as measured by both eigenvector centrality and degree. Overall, this arm of the experiment constitutes a meaningful and challenging test for the network-based targeting treatments since the extension agents were able to use valuable information not available to researchers, such as the individual's motivation to take on the role. The Benchmark treatment is similar to what the Malawi Ministry of Agriculture and other policymakers would normally do, so this is the most relevant counterfactual.¹¹

⁸ As an example, a Simple partner is a seed if the village is randomized to be a Simple village or a shadow farmer if the village is Complex, Geo, or Benchmark.

⁹ We did not ask extension workers to name the seed farmers they would choose and then ask them to train other seeds, since we thought it would lead to high noncompliance.

¹⁰ Eigenvector centrality is weighted sum of connections, where each connection's weight is determined by its own eigenvector centrality (like Google page-rank).

¹¹ Normally the Ministry only trains one "Lead Farmer" per village, not two. In most villages, the Lead Farmer will already be established, except for villages in which there hasn't been an extension officer assigned to the village for a long time. The extension agents would have had to select a second seed farmer in Benchmark villages due to the experiment.

C. Agricultural Technologies

In this section, we describe the two technologies introduced to seed farmers and in online Appendix Section A.2 we analyze data on crop yields to give further insights into the benefits of the technologies.

Pit Planting.—Maize farmers in Malawi traditionally plant seeds in either flat land or after preparing ridges. Ridging has been shown to deplete soil fertility and decrease agricultural productivity over time (Derpsch 2003, 2004). In contrast, pit planting involves planting seeds in a shallow pit in the ground, in order to retain greater moisture for the plant in an arid environment, while minimizing soil disturbance. In our sample, pit planting was not widely practiced at baseline: 9 out of 4,004 farmers (0.22 percent) planted with pits the year prior to treatment. The technique is practiced more widely in the Sahel, and has been shown to greatly enhance maize yields both in controlled trials and in field settings in East Africa, with estimated gains of 50–113 percent in yields (Haggblade and Tembo 2003, BenYishay and Mobarak 2019). In online Appendix Section A.2, we present evidence that pit planting increased yields by 44 percent (a treatment on the treated estimate) for our trained seed farmers. The enhanced productivity is thought to derive from three mechanisms: (i) reduced tillage of topsoil, which allows nutrients to remain fixed in the soil rather than eroding, (ii) concentration of water around the plants, which aids in plant growth during poor rainfall conditions, and (iii) improved fertilizer retention.

Practicing pit planting may involve some additional costs. First, hand weeding or herbicide requirements may increase because less land is tilled, though focus groups undertaken by the authors suggest that weeding demands were actually reduced substantially relative to ridging. Second, digging pits is a labor-intensive task with large up-front costs. However, land preparation becomes easier over time, since pits should be excavated in the same places each year, and estimates suggest that land preparation time falls by 50 percent within 5 years (Haggblade and Tembo 2003). BenYishay and Mobarak (2019) find that in Malawi, labor time decreases while the change in other input costs are negligible in comparison. Labor costs are minimized when pit planting is used on flat land.

Crop Residue Management.—Seed farmers were also trained in crop residue management (CRM), a set of farming practices which largely focus on retention of crop residues in fields for use as mulch. Alternative practices commonly used by farmers include burning the crop residues in the fields and removing them for use as livestock feed and compost. The trainings emphasized the value of retaining crop residues as mulch to protect topsoil, reduce erosion, limit weed growth, and improve soil nutrient content and water retention. There is little experimental evidence on the impacts of CRM on soil fertility, water retention, and yields in similar settings.

D. Seed Farmers: Descriptive Statistics, Training, and Take-Up

Extension agents chose the seed farmers in the Benchmark villages, and the researchers chose the seeds in the remaining treatment villages. We already discussed in Table 1 how central the seeds are in different treatments. Online Appendix

Table A2 provides some summary statistics describing how the chosen seeds differ in terms of farm size and a wealth index.¹² The most striking pattern is that the farmers selected as seeds under the geographic treatment are significantly poorer than other seeds. This is because many households live on one of their plots in Malawi. Households who are geographically close to lots of people will mechanically have less land, and these households tend to be poorer overall.

We observe that there are more households connected to both seeds in Complex villages than in other treatment arms. A total of 35 percent of our random household sample has a connection to a Simple partner, and 6 percent are connected to both Simple partners. By contrast, 18 percent of households are connected to two Complex partners. For the Geo-based partner, 10 percent of households are connected to two Geo partners. Online Appendix Table A3 displays the distribution of how far, in social distance, households are from the partner farmers in the different treatment arms.

In addition to the names of the two seed farmers, we provided extension agents with replacement names in all non-Benchmark villages in case either of the chosen seeds refused to participate in the training.¹³ Refusal was uncommon: extension agents trained 93 percent of the selected seeds or their spouses. We conduct intent-to-treat analysis using the original seed assignment.

The seed farmers received a small in-kind gift (valued at US\$8) if they themselves adopted pit planting in the first year. There was no gift or incentive provided on the basis of others' adoption in the village or the seeds' own adoption in subsequent years. Online Appendix Table A4 demonstrates that the training (and incentive) was effective at inducing adoption, but not perfectly. Seed farmers, relative to the shadow farmers, are more likely to know how to do pit planting and more likely to adopt pit planting during the first agricultural year. Note, 30 percent of seed farmers adopted pit planting during year 1, compared to 5 percent of shadow farmers ($p < 0.01$). Moreover, the adoption rate among seed farmers is the same across all treatment arms: Complex, Simple, Geo, and Benchmark.

Knowledge and adoption rates of pit planting increase among the shadow farmers over time. Knowledge of pit planting among the seeds is declining slightly between year 1 and years 2 and 3, but there remains a significant knowledge gap between seed and shadow farmers even in year 3. Adoption remains more or less constant among seed farmers. Online Appendix Section A.3 and the notes to online Appendix Table A4 provide the details on the econometric specification used for these results. Seed farmers are also more likely to adopt crop residue management (CRM) in year 1. However, by year 2 there is no longer a meaningful gap in the CRM adoption rate, and in fact the adoption rate among shadow farmers is declining over time. Given this pattern, and the fact that CRM was not a "new" technology in this area, we focus our analysis on the adoption of pit planting. We include CRM adoption results in online Appendix Table A6.

¹²Table 1 is not demonstrating balance in the randomization of villages across treatment arms. Note that there are only 100 Benchmark farmers since we never observe shadow Benchmark farmers.

¹³As the technologies themselves were new, the extension agents were themselves trained by staff from the Ministry's Department of Land Conservation.

II. Data

After training the seed farmers, we collected up to three rounds of household survey data. Online Appendix Figure A1 shows the timeline of these data collection activities. We describe each major data source in turn.

Social Network Census Data.—Targeting based on different network characteristics requires relatively complete information on network relationships within the village (Chandrasekhar and Lewis 2016). More than 80 percent of households in every sample village participated in the census.¹⁴

The main focus of the social network census was to elicit the names of people each respondent consults when making agricultural decisions. General information on household composition, socioeconomic characteristics of the household, general agriculture information, and work group membership was also collected. Agricultural contacts were solicited through several prompts.¹⁵ These responses were matched to the village listing to identify links. Individuals are considered linked if either party named each other (undirected graph), and all individuals within a household are considered linked.

Sample Household Survey Data.—We collected survey data on farming techniques, input use, yields, assets, and other characteristics for a sample of approximately 5,600 households in the 200 sample villages. We attempted to survey all seed and shadow farmers in each village, as well as a random sample of 24 other individuals, for a total of about 30 households in each village.¹⁶ In villages with fewer than 30 households, all households were surveyed. Three survey rounds were conducted in Machinga and Mwanza in 2011, 2012, and 2013, and two survey rounds were conducted in Nkhotakota in 2012 and 2013.¹⁷ The first round asked about agricultural production in the preceding year, thus capturing some baseline characteristics, as well as current knowledge of the technologies, which could reflect the effects of training. Since the data were collected at the start of a given agricultural season, but after land preparation was complete, we observe three adoption decisions for pit planting for farmers in Mwanza and Machinga, and two decisions for farmers in Nkhotakota. Since crop residue management (CRM) decisions are made at the end

¹⁴We interviewed at least one household member from 89.1 percent of households in Nkhotakota, 81.4 percent in Mwanza, and 88.6 percent in Machinga. We interviewed both a man and a woman in about 30 percent of households.

¹⁵We first asked in general terms about farmers with whom they discuss agriculture. To probe more deeply, we also asked them to recall over the last five years if they had (i) changed planting practices; (ii) tried a new variety of seed, for any crop; (iii) tried a new way of composting; (iv) changed the amount of fertilizer being used for any crop; (v) tried a new crop, such as paprika, tobacco, soya, cotton, or sugar cane; or (vi) started using any other new agricultural technology. If they responded affirmatively, we asked respondents to name individuals they knew had previously used the technique in the past and whether they had consulted these individuals. Finally, we asked them if they discussed farming with any relatives, fellow church or mosque members, or farmers whose fields they pass by on a regular basis, or if there are any others with whom they jointly perform farming activities. We also elicited their close friends and contacts with whom they share food, though we did not include these contacts as agricultural connections for the purposes of our network mapping.

¹⁶In Simple, Complex, and Geo villages there were 6 (2×3) seed and shadow farmers to interview, while in Benchmark villages there were 8 (2×4) seeds and shadows. Recall we do not observe Benchmark farmers in Simple, Complex, and Geo villages.

¹⁷Unanticipated delays in project funding required us to start training of extension agents and seed farmers in Nkhotakota in 2012 instead of 2011 as we did in Mwanza and Machinga.

of an agricultural season after harvest, we observe CRM decisions for two agricultural seasons in Mwanza and Machinga, and one in Nkhotakota.

Randomization and Balance.—Randomization was stratified by district, and implemented using a re-randomization procedure which checked balance on three village-level covariates.¹⁸ Online Appendix Table A5 shows how observable baseline characteristics from the social network census vary with the treatment status of the village. The table also shows *p*-values from the joint test of all treatment groups. The table notes provide details on the specification used. Few differences across treatment groups are statistically significant. Overall, the joint test reveals no differences for 10 out of 12 variables. Farm size is the most concerning: farmers in the Benchmark villages have larger farm sizes on average than farmers in Simple and Complex villages, and the joint test across the network treatment variables is significant at the 10 percent level. Additional analysis available from the authors controls for this variable in all specifications and finds that all results are robust to this control.

III. Average Treatment Effects on Diffusion

In this section, we report experimental results on village-level outcomes across the four treatment arms.

A. The Advent of Diffusion

We focus on the advent of diffusion in our sample villages as a key outcome. While the speed of diffusion may matter in some settings, we think that a key policy goal is to have diffusion start in as many villages as possible. If there is no diffusion in a village after 3 years, it is likely that the technology will never be widely adopted.

Therefore, we first focus on “any adoption” as an indicator for villages which have at least one household (other than the seeds) that adopted pit planting. Our village-level regression is as follows:

$$Y_v = \alpha + \beta_1 \text{Complex}_v + \beta_2 \text{Simple}_v + \beta_3 \text{Geo}_v + X_v\gamma + \varepsilon_v,$$

where $X\gamma$ are variables used in the re-randomization routine, specified in the table notes, and district fixed effects. The results are reported in Table 2. First note that in year 2, we observe the start of the diffusion process in only 42 percent of Benchmark villages. This increases in year 3 to a modest 54 percent. This is evidence that this is an environment where igniting diffusion is challenging. The first two columns of Table 2 show that the propensity for “any adoption” in year 2 is statistically significantly larger in villages where both seeds were highly central (Complex diffusion treatment) relative to Benchmark villages. The 25 percentage point gap is large relative to the “any adoption” rate of 42 percent in our Benchmark villages. The “any

¹⁸The three variables include: percent of village using compost at baseline, percent of village using fertilizer at baseline, and percent of village using pit planting at baseline. We control for these variables in the analysis.

TABLE 2—VILLAGE-LEVEL REGRESSIONS OF ADOPTION OUTCOMES ACROSS TREATMENT ARMS

	Any non-seed adopters		Adoption rate	
	(1)	(2)	(3)	(4)
Complex diffusion treatment	0.252 (0.093)	0.304 (0.101)	0.036 (0.016)	0.036 (0.026)
Simple diffusion treatment	0.155 (0.100)	0.189 (0.111)	0.036 (0.017)	0.006 (0.022)
Geographic treatment	0.107 (0.096)	0.188 (0.110)	0.038 (0.027)	0.013 (0.034)
Year	2	3	2	3
Observations	200	141	200	141
Mean of Benchmark treatment (omitted category)	0.420	0.543	0.038	0.075
SD of Benchmark	0.499	0.505	0.073	0.109
<i>p-values for equality in coefficients</i>				
Simple = Complex	0.300	0.240	0.981	0.173
Complex = Geo	0.102	0.220	0.937	0.491
Simple = Geo	0.623	0.990	0.950	0.783

Notes: The reference group is the Benchmark treatment. The sample for year 3 (columns 2 and 4) excludes Nkhotakota district. The *Any non-seed adopters* indicator in columns 1–2 excludes seed farmers. The adoption rate in columns 3–4 include all randomly sampled farmers, excluding seed and shadow farmers. All columns include controls used in the re-randomization routine (percent of village using compost at baseline; percent village using fertilizer at baseline; percent of village using pit planting at baseline); village size and its square; and district fixed effects. Standard errors are clustered at the village level.

adoption” rate in Complex villages is also 15 percentage points larger than in Geo villages ($p = 0.10$) and 10 percentage points larger compared to villages assigned to the simple diffusion treatment ($p = 0.30$). In year 3, Simple, Complex and Geo villages all attain a statistically higher rate of “any adoption” than Benchmark villages. Here, 85 percent of Complex villages had at least one non-seed adopter, compared to 73 percent of Simple and Geo villages and 54 percent of Benchmark villages.

B. Adoption Rates across Treatment Arms

We also look at the speed of diffusion, captured by the adoption rate. Columns 3 and 4 in Table 2 document treatment effects on the adoption rate, which is defined as the proportion of non-seed farmers who adopted pit planting in each agricultural season. Both Simple and Complex diffusion villages have higher adoption rates relative to the Benchmark in year 2. Compared to the Benchmark rate of 3.8 percent, Complex and Simple villages both experience a 3.6 percentage point higher adoption rate. We cannot reject that the adoption rates are the same in Simple, Complex, and Geo villages. The adoption rate increases across all four types of villages in year 3. The adoption rate increases in the Benchmark villages, the reference category, from 3.8 percent to 7.5 percent from years 2 to 3. With the smaller sample size of 141 villages in year 3, we cannot reject that the adoption rate is the same across all treatment types, though the point estimate on Complex remains the largest, and is equal in magnitude to the effect size observed in year 2. The adoption rate in Complex villages in year 3 is 11 percent.

C. Discussion: Why Did Targeting Central Seeds Matter?

Targeting central seeds as in the Complex treatment led to higher adoption and was particularly important for avoiding the scenario in which no farmers adopted at all. In many diffusion models, this total failure of adoption would be quite surprising: generically, nearly everyone is connected to the network, and so some diffusion should have taken place in Benchmark villages, too. Akbarpour, Malladi, and Saberi (2020) describe characteristics of diffusion processes where targeting has an advantage. First, in early stages of diffusion, targeting will speed up the adoption process. But with time, the diffusion process in Benchmark villages could catch up to Complex diffusion villages. However, the results in Section IIIA suggest that for many villages, a longer time horizon will not lead to substantially more adoption. With virtually no adoption after 3 years, it is unlikely those villages will ever have widespread adoption of pit planting.

Second, when information sharing is sufficiently infrequent, targeting may matter. We use data on conversations about pit planting that respondents had with others in the village to look directly at this explanation. Each respondent was asked questions about their relationship and conversations with the two seed farmers, randomly selected shadow farmers, and a random sample of other village residents.

Approximately 18 percent of farmers report talking about pit planting with trained seeds each year. This is a reasonably high rate of information passing, such that we would anticipate that the AMS dynamics of information eventually reaching the central farmers would be at play in a SIR-type model. In fact we also observe that many (13–14 percent of respondents) are also having conversations with shadow partners about pit planting, likely because of those very dynamics. We can provide a lower bound on how much the experiment induced additional conversations about pit planting using the random variation in the experiment itself. For example, we compare the frequency of conversations with the Complex seed farmers in Complex diffusion villages, to the frequency of conversations with Complex shadow farmers in other villages. This is a conservative, downwardly based estimate as many (and perhaps most, given how unusual pit planting was at baseline) of the conversations with Complex shadow farmers will also have occurred because of the experiment. However, this conservative estimate is sufficient to argue that it is unlikely that farmers are not talking enough to generate an adoption cascade.

Table 3 shows that the experiment indeed induced seed farmers to discuss pit planting with fellow villagers using the following econometric specification:

$$Y_{ij} = \alpha + \beta_1 \text{Trained}_j + \delta_1 \text{ComplexPartner}_j \\ + \delta_2 \text{SimplePartner}_j + \delta_3 \text{GeoPartner}_j + X\gamma + \varepsilon_v.$$

Here, Y_{ij} is an indicator for whether respondent i discussed pit planting with partner (either seed or shadow) farmer j ; Trained_j is 1 if a partner was trained in pit planting¹⁹ and 0 otherwise; ComplexPartner_j is an indicator for whether the partner j is

¹⁹This arises for complex partners in Complex diffusion villages, Simple partners in Simple diffusion villages, and Geo partners in Geo villages.

TABLE 3—CONVERSATIONS FARMERS REPORT HAVING ABOUT PIT PLANTING WITH SEED AND SHADOW PARTNERS

	Conversation about pit planting		
	(1)	(2)	(3)
Trained	0.037 (0.008)	0.050 (0.008)	0.064 (0.009)
Percent conversation with trained seed	0.179	0.181	0.190
Percent conversation with shadow partner	0.141	0.130	0.127
Observations	15,115	16,704	11,607
Year	1	2	3

Notes: The sample excludes seeds and counterfactual/shadow farmers. In our survey, we asked respondents about conversations they had with the seed farmers and randomly selected counterfactual/shadow farmers. In this table, we refer to farmers who would be seeds under the different treatments as partners, whether they are trained seeds or are shadow farmers. An observation is a respondent-partner-year pair. The following indicator variables are also included in the regressions: whether the contact that the respondent was asked about was a simple partner, complex partner or geo partner, irrespective of whether they were trained. All columns include controls used in the re-randomization routine (percent of village using compost at baseline; percent village using fertilizer at baseline; percent of village using pit planting at baseline); village size and its square; and district fixed effects. Standard errors are clustered at the village level.

a complex partner (either seed or shadow) and *SimplePartner_j* and *GeoPartner_j* are defined analogously; $X\gamma$ are variables used in the re-randomization routine, specified in the table notes, and district fixed effects; β_1 is our coefficient of interest. Since we only consider conversations with treated partners and shadow partners, whether a potential conversation partner was actually trained is random and we can interpret the effect of training on conversations as exogenous.

We find that about 5 percent (ranging from 3.7 percent in year 1 to 6.4 percent in year 3) more respondents report a conversation about pit planting with trained seeds than with untrained seeds. In online Appendix Section A.4, we suppose that only these 5 percent of conversations are attributable to the training, and find that this lower bound exceeds the conversation threshold AMS establish in which random seeding should generate an adoption cascade in simple diffusion models.

An additional possibility that AMS highlight is that targeting may be more important in a range of diffusion models outside of the class of “simple diffusion” models they consider; in the next section, we consider an important model outside of this class: the threshold diffusion model.

IV. Complex Contagion

In this section, we propose that the threshold model we used to select seeds offers a potential explanation for why targeting central seeds matters for diffusion. As AMS make clear and Jackson and Storms (2019) formalize, targeting will be advantageous relative to random seeding when diffusion is governed by a threshold model. The intuition for the importance of targeting in the threshold model is illustrated with the example network shown in Figure 1. In this thought experiment, we train two seed farmers in period 0 such that they are fully informed about a new technology. Diffusion occurs as farmers become informed in subsequent periods.

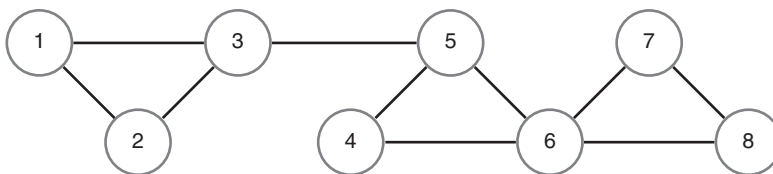


FIGURE 1. AN EXAMPLE NETWORK

Suppose that farmers in this network become fully informed of a new technology if anyone they are connected to has been fully informed. This is what we call simple contagion. In this network, the ideal seed farmers will be farmer 6 and then either farmer 1, 2, or 3. With any of these configurations, all farmers are informed in period 1. In general, quickly diffusing information about the new technology will be easy: in 70 percent of all possible seed pairings, all farmers will be fully informed by the end of the second period. Targeted seeding is not necessary in this model.

However, if farmers need to know two other farmers before they have sufficient information to be fully informed, the diffusion process looks very different. Consider seeding farmers 5 and 8. During the first period, farmer 6 will become informed. In the second round, farmers 4 and 7 are informed. The diffusion process then stops with 3 out of a possible 6 non-seed farmers informed. There are 4 seed pairings which can achieve this 50 percent adoption rate, but it is not possible to get any higher.

Crucially, without a focus on targeting, there is a good probability that there is no diffusion: in 40 percent of seed pairings, there is no diffusion whatsoever. Threshold models therefore generate the empirical result we observed: when noncentral farmers are trained, there may be no diffusion at all.

In the next subsections, we will provide a micro-foundation of the threshold model based on social learning. We then provide three pieces of empirical evidence that are consistent with the idea that complex contagion is a reason why targeting central seeds was effective in this setting.

A. A Micro-Foundation for the Threshold Model of Diffusion

Social learning is known to be important in technology adoption decisions (e.g., Griliches 1957, Conley and Udry 2010). This section demonstrates how social learning naturally micro-founds the threshold model. Our theoretical framework considers a learning environment with three characteristics. First, we suggest that adoption of a new technology takes place only when farmer beliefs about the profitability of the technology pass a critical threshold. Second, there are limited inherent benefits to learning about a technology if farmers are not ultimately persuaded to adopt it. Third, learning is costly: farmers must invest time to learn about and master a new productive technology, and revealing ignorance may subject them to social costs (e.g., Banerjee et al. 2020; Chandrasekhar, Golub, and Yang 2019).

These facts together mean that technology diffusion will be characterized by rational ignorance: farmers will be unwilling to pay learning costs in environments

where they are unlikely to update their beliefs enough to adopt the new technology. Moreover, if farmers aggregate multiple signals to update their beliefs via Bayes' rule, technology adoption will be characterized by multiple equilibria: when few are informed about the technology, few will be willing to pay learning costs and few will adopt; when many are informed, more farmers will pay learning costs and ultimately adopt.

In online Appendix Section A.5, we adapt the naïve learning model in Banerjee et al. (2016) to include small costs of learning. We model technology diffusion as a learning process with three key phases: (i) the farmer has to decide whether to acquire information, (ii) she combines the new information with her priors via Bayes' rule, and (iii) based on her revised information set, she then decides whether to adopt the new technology. We demonstrate that farmers who learn in this way follow a threshold model (Granovetter 1978; Acemoglu, Ozdaglar, and Yildiz 2011): a farmer will become informed about a new technology once at least λ of her connections become informed. Since uninformed farmers do not adopt, this means that farmers without sufficient informed connections will not adopt.

Taking the model to the data, the micro-foundation is useful for a few purposes. First, it demonstrates that agricultural learning can lead to diffusion with thresholds. In our micro-foundation, thresholds arise because farmers rationally choose not to learn when there is insufficient information in the network to change their behavior. This suggests that the learning problem could generate the results in Section IIIA because farmers need to be exposed to multiple informed agents to make an informed adoption decision. As a result, targeting central farmers is critical for diffusion: poor targeting may lead to no diffusion at all. Second, we can characterize the learning problems which lead to higher thresholds: thresholds are higher when the expected benefits are lower, or when signals are noisier. We therefore learn that seeding matters more in contexts where (in expectation) the benefits of adoption are relatively low; or in cases where a given signal is quite likely to be noisy. We return to this prediction in Section IVC.

B. Complex Contagion Model Simulations Compared to Empirical Results

There are three main pieces of evidence that suggest that complex contagion may have led to higher diffusion in the villages in which both seeds had high centrality. First, a key consequence of not targeting the right seeds in an environment where a sizable fraction of agents have a threshold above 1 is that the diffusion process can be completely stalled. We will discuss this evidence in this subsection. Second, we show heterogeneous treatment effects to argue that the complex diffusion treatment was particularly effective in exactly the contexts in which we would anticipate the treatment to be effective. And finally, we analyze individual-level data to show that farmers who were directly connected to two seeds as opposed to just one seed are most likely to adopt pit planting.

Table 2 already demonstrated that complex diffusion led to a higher rate of “any non-seed adoption.” Figure 2 presents the same evidence but side by side with what our simulations predicted. The left part of Figure 2 shows the *predicted* fraction of villages with “any adoption” from simulating the model for all sample villages when

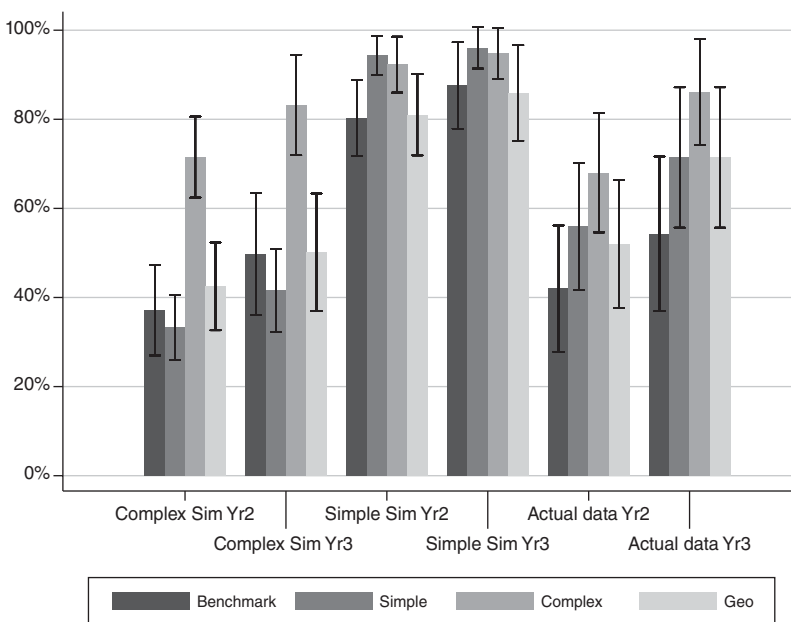


FIGURE 2. PERCENT OF VILLAGES WHERE AT LEAST SOME NON-SEEDS ADOPTED IN DATA AND SIMULATIONS

$\lambda = 1$ (Simple contagion) and $\lambda = 2$ (Complex contagion).²⁰ Since the goal is to compare these simulations to the actual data, we design the simulations to reflect the fact that we only observe a random sample of households in these villages.²¹ The right part of Figure 2 shows the empirical counterpart: “any adoption” rates in the data in years 2 and 3.

When the threshold is set to $\lambda = 1$, diffusion is predicted to be widespread. In year 2, 85 percent of villages where Geo and Benchmark partners were trained are predicted to have some sampled diffusion, and that rate goes up to 94 percent with Simple and Complex partners. The predicted rates of “any diffusion” are even higher in year 3.

The risk of no diffusion increases if the diffusion process is characterized by complex contagion. In that case, the model predicts that more than half of the villages assigned Simple, Geo, or Benchmark partners will not see any sampled diffusion at all in year 2. In contrast, when Complex seeds are trained, 70 percent of villages are predicted to experience some diffusion in year 2.

²⁰These simulations exclude 12 villages where at least one of the extension worker chosen seeds (Benchmark) was not observed in our social network census. This occurred because the spatial boundaries of villages are not always clearly delineated, particularly in Nkhotakota.

²¹The simulations use the full social network to predict becoming informed, measured here through adoption. We then sample from the full network to better mimic our data. In the model, the rate of any adoption is identical in years 2 and 3. If there was no adoption by year 2, there is no way there will be any additional adoption taking place in year 3. The sampling process, however, generates the increase over time observed in the figure. If the rate of adoption is low, as is empirically the case, then a random sample may miss all adopters. As the number of adopters increases over time, the random sample is more likely to pick up an adopter and hence the rate of any adoption increases over time in the figure.

Comparing the theoretical simulations to the data on the right side of Figure 2 shows that the data are more consistent with the patterns generated by a complex (rather than simple) learning environment in three distinct ways. First, simple contagion simulations suggest that we should observe a much higher fraction of villages with some adoption than is true in the data. Second, simple contagion predicts that the “any adoption” outcome should not be very sensitive to the identity of the seed farmer who is initially trained. In contrast, the identity of the seed farmer dramatically alters this outcome in the data. Finally, the complex contagion simulations predict that the Complex partners will maximize the fraction of villages with some adoption, which we observe in the data.

C. Heterogeneity Analysis

The micro-foundation of the threshold model suggests that targeting Complex diffusion seeds will be particularly effective in contexts in which the information about pit planting will be most valuable. We use two different approaches to identify groups of such farmers. First, the Ministry of Agriculture recommends pit planting only for flat land, and labor costs of pit planting are lower on flat land.²² Focus group discussions in our sample villages confirmed that villagers thought pit planting was more suitable for flat rather than sloped land. We therefore expect farmers who own flat land will be most interested in information about pit planting. The second heterogeneity test we do exploits variation in knowledge about pit planting at baseline. While pit planting is in general a new technology in Malawi, there is heterogeneity across villages in how novel it is. In the median village, 4.3 percent of farmers reported having ever tried pit planting at baseline while 0.2 percent were currently practicing pit planting across all villages.

Table 4 explores the heterogeneity in treatment effects across these two dimensions by interacting the randomized treatments with an indicator for “Farmer likely to receive a Good Signal.” This *Good Signal* variable is first defined as the farmer having flat land in columns 1 and 2, and then redefined as “Village with lower-than-median familiarity with the technology at baseline” in columns 3 and 4. *Bad Signal* refers to the converse of these characteristics. The equation estimated

$$\begin{aligned}
 y_{ivt} = & \beta_0 + \beta_1 Simple_v \times BadSignal + \beta_2 Complex_v \times BadSignal \\
 & + \beta_3 Geo_v \times BadSignal + \beta_4 GoodSignal + \beta_5 Simple_v \times GoodSignal \\
 & + \beta_6 Complex_v \times GoodSignal + \beta_7 Geo_v \times GoodSignal + \delta X_v + \epsilon_{ivt}.
 \end{aligned}$$

The reference group comprises of farmers who are likely to receive a bad signal in Benchmark villages. Our hypothesis is that among those who receive a positive signal, we will observe more diffusion in Complex villages if the true model is Complex.

²² Pit planting is possible on land with some slope, but in those cases, the pits need to be constructed differently, and our extension workers were not trained on that technique.

TABLE 4—HETEROGENEITY IN FARMER-LEVEL ADOPTION DECISIONS ACROSS TREATMENT ARMS

	(1)	(2)	(3)	(4)
Bad Signal × Complex	0.006 (0.024)	−0.027 (0.036)	0.013 (0.015)	−0.045 (0.033)
Bad Signal × Simple	−0.008 (0.024)	−0.036 (0.037)	0.019 (0.017)	−0.008 (0.034)
Bad Signal × Geo	0.002 (0.031)	−0.068 (0.031)	0.031 (0.035)	−0.054 (0.032)
Good Signal	−0.037 (0.017)	−0.062 (0.024)	−0.007 (0.022)	−0.064 (0.038)
Good Signal × Complex	0.059 (0.018)	0.067 (0.025)	0.054 (0.024)	0.083 (0.030)
Good Signal × Simple	0.064 (0.021)	0.029 (0.020)	0.054 (0.029)	0.021 (0.020)
Good Signal × Geo	0.042 (0.020)	0.022 (0.023)	0.026 (0.022)	0.031 (0.029)
Good Signal type	Flat land	Flat land	Unfamiliar tech	Unfamiliar tech
Year	2	3	2	3
Observations	3,546	2,645	3,954	3,023
Mean of Bad Signal in Benchmark treatment (omitted category)	0.066	0.123	0.046	0.104
SD	0.248	0.33	0.21	0.305
<i>p-values for equality in coefficients</i>				
Simple, good = Complex, good	0.828	0.113	0.986	0.032
Complex, good = Geo, good	0.482	0.103	0.297	0.138
Simple, good = Geo, good	0.364	0.755	0.351	0.680

Notes: The reference group is *Bad Signal* recipients in the Benchmark treatment. In columns 1 and 2, households with any flat land are those who have *Good Signal* = 1 and those with all sloped land have *Good Signal* = 0. In columns 3 and 4, households in villages where less than 4.32 percent (the median) of households ever tried pit planting at baseline are those who have *Good Signal* = 1. Sample for year 3 (columns 2 and 4) excludes Nkhotakota district. All columns include controls used in the re-randomization routine (percent of village using compost at baseline; percent village using fertilizer at baseline; percent of village using pit planting at baseline); village size and its square; and district fixed effects. Standard errors are clustered at the village level.

Columns 1 and 2 show that adoption in year 2 is higher for farmers who have flat land in Simple, Complex, and Geo villages compared to farmers with flat land in Benchmark villages. In year 3, we see that Complex villages continue to have a larger adoption rate than Benchmark villages for farmers with flat land. Columns 3 and 4 show that the Complex treatment performs best in villages where the technology was relatively novel. In this subsample, the adoption rate is statistically significantly higher in Complex diffusion treatment villages compared to both the Simple and the Benchmark treatments in year 3.

To summarize, these heterogeneity tests indicate that targeting central seeds is most effective precisely in the types of villages and for the types of farmers where information was most valuable, as the theoretical model helped us predict.

D. Knowledge and Adoption of Farmers by Social Distance to Seeds

In this subsection, we provide more direct evidence in line with the Complex Contagion model. We look at knowledge of pit planting and adoption decisions by individuals, as a function of how many seeds they are connected to. If thresholds are

larger than one, those with connections to 2 seeds should be the most likely to adopt pit planting. Our identification strategy is summarized by the equation:

$$Y_{iv} = \alpha + \beta_1 1TSeeds_{iv} + \beta_2 2TSeeds_{iv} + \beta_3 1Simple_{iv} + \beta_4 2Simple_{iv} \\ + \beta_5 1Complex_{iv} + \beta_6 2Complex_{iv} + \beta_7 1Geo_{iv} + \beta_8 2Geo_{iv} + \theta_v + \varepsilon_{iv},$$

where *1TSeeds* is an indicator for the respondent being directly connected to exactly one seed farmer, and *2TSeeds* indicates the respondent was directly connected to two seed farmers. Seeds and shadows are removed from the analysis. Since network position is endogenous, we also control for whether an individual is connected to one or two Simple, Complex, or Geo (actual or shadow) partners, but these coefficients are not displayed in the table. Identification therefore comes from variation in the experiment. As an example, we can compare two farmers who are both connected to two Simple partners, but where one farmer is in a village randomly assigned to the Simple treatment and his friend is trained as the seed, while the other farmer's friend was not trained.

In the theoretical model, individuals have to become informed prior to adopting. As an empirical matter, it is unclear what level of knowledge is associated with "being informed" as used in the model. In Table 5, we therefore consider three variables which represent increasing levels of information: whether the respondent has heard of pit planting; whether the respondent knows how to implement pit planting; and whether the respondent adopted pit planting (which implies not only knowledge but also that the signals that the respondent received were sufficiently positive). In year 1, the training led to more information transmission to those directly connected to seeds. In particular, those who have a direct connection to both seed farmers had the most knowledge. This is true for both measures of "knowledge": whether the respondent had heard of pit planting and whether they reported being capable of implementing it. Respondents with two connections are 8.4 percentage points more likely to have heard of pit planting than those with no connection to a seed. This represents a 33 percent increase in knowledge relative to the mean familiarity among unconnected individuals. This effect is also statistically significantly different from the effect of being connected to one seed ($p = 0.02$). They are also 6.2 percentage points more likely to report knowing how to pit plant, a 108 percent increase over unconnected individuals and again significantly different from the effect of being connected to one seed ($p = 0.072$). These knowledge effects are suggestive, but not conclusive, of a complex contagion process ($l = 2$) rather than simple contagion. The increased awareness of pit planting and knowledge of pit planting among households connected to two seeds persists into year 2 (columns 2 and 5), and two connections is again significantly more advantageous than one connection ($p = 0.04$ and 0.095 , respectively).

In Table 6, we look at adoption of pit planting. We see no effect on adoption in the first year (column 1) among individuals directly connected to either one or two seeds. However, we do observe an adoption effect in year 2. This temporal pattern of results is consistent with the set-up of our theoretical model: individuals become informed in year 1 and then some choose to adopt in year 2. Column 2 shows that households with two connections to trained seeds are 3.9 percentage points more

TABLE 5—DIFFUSION WITHIN THE VILLAGE: KNOWLEDGE

	Heard of pit planting			Knows how to pit planting		
	(1)	(2)	(3)	(4)	(5)	(6)
Connected to 1 seed	0.002 (0.024)	0.030 (0.022)	0.016 (0.029)	0.017 (0.016)	0.021 (0.017)	-0.031 (0.023)
Connected to 2 seeds	0.084 (0.038)	0.124 (0.040)	0.064 (0.064)	0.062 (0.028)	0.068 (0.029)	0.110 (0.051)
Within path length 2 of at least one seed	-0.018 (0.028)	0.016 (0.027)	0.067 (0.042)	0.005 (0.018)	0.022 (0.021)	0.028 (0.028)
Year	1	2	3	1	2	3
Observations	4,155	4,532	3,103	4,155	4,532	3,103
Mean of reference group (no connection to any seed)	0.223	0.286	0.391	0.057	0.095	0.147
SD of reference group	0.416	0.452	0.488	0.232	0.293	0.355
<i>p</i> -value for 2 connections = 1 connection	0.018	0.013	0.442	0.072	0.091	0.004

Notes: Sample excludes seed and shadow farmers. The reference group is comprised of individuals with no direct or 2-path-length connections to a seed farmer. Only connections to simple, complex, and geo seed farmers are considered (no connections to Benchmark farmers included). The dependent variable in columns 1–3 is an indicator for whether the respondent reported being aware of a plot preparation method other than ridging and then subsequently indicated awareness of pit planting in particular. In columns 4–6, the dependent variable is an indicator for whether the farmer reported knowing how to implement pit planting. In all columns, additional controls include indicators for the respondent being connected to: one Simple partner, two Simple partners, one Complex partner, two Complex partners, one Geo partner, two Geo partners, within 2 path length of a Simple partner, within 2 path length of a Complex Partner, and within 2 path length of the geo partner. Also included are village fixed effects. Standard errors are clustered at the village level.

TABLE 6—DIFFUSION WITHIN THE VILLAGE: ADOPTION

	Adopts pit planting		
	(1)	(2)	(3)
Connected to 1 seed	0.008 (0.011)	0.012 (0.015)	0.004 (0.017)
Connected to 2 seeds	0.016 (0.014)	0.039 (0.019)	0.014 (0.035)
Within path length 2 of at least one seed	0.013 (0.008)	0.022 (0.013)	0.037 (0.021)
Year	1	2	3
Observations	4,203	3,931	2,998
Mean of reference group (no connection to any seed)	0.013	0.044	0.043
SD of reference group	0.113	0.206	0.203
<i>p</i> -value for 2 connections = 1 connection	0.522	0.164	0.760

Notes: See notes in Table 5 for details on the specification. The dependent variable in columns 1–3 is an indicator for the household having adopted pit planting in that year.

likely to adopt in the second year than those with no connections, which represents a 90 percent increase in adoption propensity. Though the point estimate of the effect of two connections is considerably larger than the effect of a connection to one seed (3.9 pp compared to 1.2 pp), we cannot statistically reject that households with a connection to only one treated seed adopt less frequently ($p = 0.16$). We also observe that individuals who are within path length 2 of at least one seed (that is, a friend of a friend) are 2.2 percentage points more likely to adopt.

The predictions of the model for which individuals learn about pit planting are weakened as time passes and knowledge diffuses through the network. In all three of the dependent variables in Tables 5 and 6, this diffusion can be observed through large increases in knowledge and adoption over time in our reference category: individuals with no direct connections to a seed. Among this group awareness increases from 22 percent to 39 percent from year 1 to 3, while “knowing how” to pit plant increases from 6 percent to 15 percent and adoption increases from 1 percent to 4 percent. In principle, this diffusion should reduce power on our exogenous variation, as the number of connections to informed individuals becomes less correlated with the number of signals available to farmers. In practice, by year 3 we still see significance on the effects of two direct connections on one of our two knowledge variables (“knowing how” to pit plant, column 6 of Table 5), but we no longer see significant differences from direct connections in adoption or awareness of pit planting. Consistent with the hypothesis that this loss in precision is due to diffusion in the network, we see that adoption increases among those at moderate distance to the seeds in year 3: column 3 of Table 6 shows that households within path length 2 are more likely (3.7 pp) to have adopted over those who are socially more distant.²³

In summary, analysis using individual-level data demonstrates that individuals who are initially close to the trained seeds are more likely to adopt than individuals with no direct connections, as one would expect if the experiment is inducing social network-based diffusion. The data also suggest that having two direct connections, and not just one, is important for diffusion. This is further evidence consistent with the complex contagion model: farmers may need to know multiple informed connections before becoming informed, and then subsequently adopting, themselves.

V. Cost-Effective, Policy-Relevant Alternatives to Data-Intensive Targeting Methods

Our experiment was designed to be a proof of concept. We showed that targeting multiple highly central farmers improves technology diffusion, but eliciting the social network map to achieve these gains is expensive. Our geography-based treatment arm was an attempt to assess how much of the diffusion benefit derived from applying network theory could be achieved without having to resort to expensive data collection methods (since each household’s physical location is much easier to observe than network relationships). This specific approach was not an unqualified success. Online Appendix Table A2 showed that Geo seeds tended to have less land and were therefore poorer. Therefore, while the idea of using geography as a proxy for one’s network may be intuitive, the implications of geographic centrality may be context-specific, and inappropriate as a network-based targeting proxy in some cases.

Combining our experimental results with research on other inexpensive procedures to identify the optimal seeds under complex contagion theory would make network-based targeting more policy relevant and scalable. A few recent papers have suggested promising, less expensive methods for inferring network characteristics.

²³This is a lower power test of the model than the direct connections test as it is imperfectly correlated with the number of informed, indirect connections to seeds (which is unobserved). We do not see a significant effect of this variable on knowledge outcomes, though coefficients are positive.

Banerjee et al. (2019) suggests that despite the implicit challenges in learning about network structure, the simple question of “if we want to spread information about a new loan product to everyone in your village, to whom do you suggest we speak?” is successful in identifying individuals with high eigenvector centrality and diffusion centrality, who ultimately improve the diffusion process. Breza et al. (2020) suggest that aggregate relational data collected from a smaller sample combined with a census can yield accurate estimates of network characteristics. Mobile phones may also be a way to inexpensively identify highly central individuals (Björkegren 2019; Blumenstock, Chi, and Tan 2019).

While we cannot test the viability of these approaches with our data, we can explore via simulations some alternate strategies that extension officers could use to identify useful partners. We suppose that an extension agent enters a village and randomly selects a small number of farmers to interview, and only asks one question from our social network census: “Do you discuss agriculture frequently with anyone in the village? What is the name of the person you speak with about agriculture frequently?” The response to this question generates a small list of names. The extension agent can then use the responses to select any follow-up interviews. Using simulations, we predict that strategies which leverage the highest degree respondent from the random sample can approach the performance of the optimal targeting. More specifically, we can achieve 73 percent of the optimal adoption rate with just 2 total interviews and 84–90 percent of the targeting gains with around 7 interviews.²⁴

VI. Concluding Remarks

This paper provides evidence that diffusion of a new technology is accelerated by targeting information to central nodes within a social network. In a field experiment conducted in collaboration with the Ministry of Agriculture in Malawi, we selected farmers at different positions in the village network, leveraging threshold theory to suggest useful partners under different diffusion mechanisms. We found that farmers were most likely to adopt pit planting in villages where the two trained seed farmers were centrally located within their villages’ social network. These partners were chosen to optimize diffusion under complex contagion: when thresholds for diffusion were larger than 1.

Because two central partners may be optimal under several diffusion models, we also explore whether the underlying diffusion process is well characterized by complex contagion. We present multiple pieces of evidence consistent with this mechanism. In particular, we demonstrate that a total failure of diffusion occurs frequently in villages where experts selected the seed farmers. High thresholds can generate this risk. Moreover, farmers who are connected to two seed farmers are also most likely to adopt pit planting in the second year of the experiment. This is consistent with the fact that under complex contagion, multiple connections to seeds are needed before farmers adopt.

The methodological approach in this paper is not directly scalable for policy because of the high costs of collecting network data. But there is very promising

²⁴See online Appendix Section A.6 for more details and alternative targeting strategies.

work in the literature on ways to cost-effectively identify central individuals within social networks (Banerjee et al. 2019). Our simulations also suggest that with only about 7 interviews per village, it is possible to identify individuals who can trigger the diffusion process. There are also additional options available to identify central nodes within a network depending on the context, including new approaches such as cell phone data.

Our paper also suggests a direction for future research. We provide evidence that agricultural technologies need to be seeded with multiple, central individuals to encourage adoption; this and other evidence in this paper is inconsistent with “simple” diffusion models. In contrast, the evidence in this paper is consistent with models where diffusion requires a concentration of information, such as complex contagion. Further research is needed to understand if farmers often face high thresholds to adoption. Our micro-founded diffusion model suggests a key dimension to consider when assessing if contagion is likely to be simple or complex: the noise of the signal. Rosenzweig and Udry (2020) highlight the importance of aggregate stochastic shocks in distinguishing the returns to agricultural investment, microenterprise investment, and human capital from large-scale survey data. Farmers, entrepreneurs, and parents likely have access to far fewer data points than these large-scale surveys when they attempt to infer the returns to investments and schooling, which, together with our model, may suggest that high thresholds bind for a number of problems of interest to economists. However, in contexts in which agents are learning about concepts that are less noisy than returns, say the availability of microfinance, how to enroll in social protection programs, or whether a firm is hiring, simple contagion may be the right model. Characterizing which productive investments should diffuse easily through social networks, and which need extensive and targeted diffusion, is crucial but beyond the scope of this paper.

REFERENCES

- Acemoglu, Daron, Asuman Ozdaglar, and Ercan Yildiz.** 2011. “Diffusion of Innovations in Social Networks.” Paper presented at 2011 50th IEEE Conference on Decision and Control, Orlando, FL.
- Akbarpour, Mohammad, Suraj Malladi, and Amin Saberi.** 2020. “Just a Few Seeds More: Value of Targeting for Diffusion in Networks.” Unpublished.
- Alvarez, Fernando E., Francisco J. Buera, and Robert E. Lucas, Jr.** 2013. “Idea Flows, Economic Growth, and Trade.” NBER Working Paper 19667.
- Anderson, Jock R., and Gershon Feder.** 2007. “Agricultural Extension.” In *Handbook of Agricultural Economics*, Vol. 3, edited by Robert Evenson and Prabhu Pingali, 2343–78. Amsterdam: Elsevier.
- Bandiera, Oriana, and Imran Rasul.** 2006. “Social Networks and Technology Adoption in Northern Mozambique.” *Economic Journal* 116 (514): 869–902.
- Banerjee, Abhijit, Emily Breza, Arun G. Chandrasekhar, and Benjamin Golub.** 2020. “When Less Is More: Experimental Evidence on Information Delivery During India’s Demonetization.” NBER Working Paper 24679.
- Banerjee, Abhijit, Emily Breza, Arun G. Chandrasekhar, and Markus Mobius.** 2016. “Naive Learning with Uniformed Agents.” Unpublished.
- Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson.** 2013. “The Diffusion of Microfinance.” *Science* 341 (6144): 1236498.
- Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson.** 2019. “Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials.” *Review of Economic Studies* 86 (6): 2453–90.
- Beaman, Lori A.** 2012. “Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S.” *Review of Economic Studies* 79 (1): 128–61.

- Beaman, Lori, Ariel BenYishay, Jeremy Magruder, and Ahmed Mushfiq Mobarak.** 2018. "Can Network Theory-Based Targeting Increase Technology Adoption?" AEA RCT Registry. August 06.
- Beaman, Lori, Ariel BenYishay, Jeremy Magruder, and Ahmed Mushfiq Mobarak.** 2021. "Replication Data for: Can Network Theory-Based Targeting Increase Technology Adoption?" American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E130605V1>.
- BenYishay, Ariel, and Ahmed Mushfiq Mobarak.** 2019. "Social Learning and Incentives for Experimentation and Communication." *Review of Economic Studies* 86 (3): 976–1009.
- Bennett, Magdalena, and Peter Bergman.** 2020. "Better Together? Social Networks in Truancy and the Targeting of Treatment." *Journal of Labor Economics* 39 (1): 1–36.
- Björkegren, Daniel.** 2019. "The Adoption of Network Goods: Evidence from the Spread of Mobile Phones in Rwanda." *Review of Economic Studies* 86 (3): 1033–60.
- Blumenstock, Joshua, Guanghua Chi, and Xu Tan.** 2019. "Migration and the Value of Social Networks." Unpublished.
- Breza, Emily, Arun G. Chandrasekhar, Tyler H. McCormick, and Mengjie Pan.** 2020. "Using Aggregated Relational Data to Feasibly Identify Network Structure without Network Data." *American Economic Review* 110 (8): 2454–84.
- Burlig, Fiona, and Andrew Stephens.** 2019. "Reap What Your Friends Sow: Social Networks and Technology Adoption." Unpublished.
- 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): 1273–301.
- Centola, Damon, and Michael Macy.** 2007. "Complex Contagions and the Weakness of Long Ties." *American Journal of Sociology* 113 (3): 702–34.
- Chandrasekhar, Arun G., Benjamin Golub, and He Yang.** 2019. "Signaling, Shame, and Silence in Social Learning." Unpublished.
- Chandrasekhar, Arun G., and Randall Lewis.** 2016. "Econometrics of Sampled Networks." Unpublished.
- Coleman, James, Elihu Katz, and Herbert Menzel.** 1957. "The Diffusion of an Innovation Among Physicians." *Sociometry* 20 (4): 253–70.
- Conley, Timothy G., and Christopher R. Udry.** 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100 (1): 35–69.
- Derpsch, Rolf.** 2003. "Conservation Tillage, No-Tillage and Related Technologies." In *Conservation Agriculture: Environment, Farmers Experiences, Innovations, Socio-economy, Policy*, edited by Luis García-Torres, José Benites, Armando Martínez-Vilela, and Antonio Holgado-Cabrera, 181–90. Dordrecht: Springer.
- Derpsch, Rolf.** 2004. "History of Crop Production, with and without Tillage." *Leading Edge* 3 (1): 150–54.
- Duflo, Esther, and Emmanuel Saez.** 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 118 (3): 815–42.
- Flodgren, Gerd, Elena Parmelli, Gaby Doumit, Melina Gattellari, Jeremy Grimshaw, and Mary Ann O'Brien.** 2007. "Local Opinion Leaders: Effects on Professional Practice and Health Care Outcomes." *Cochrane Database Systematic Review* 24 (1): CD000125.
- 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): 1176–209.
- Granovetter, Mark.** 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83 (6): 1420–43.
- Griliches, Zvi.** 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica* 25 (4): 501–22.
- Haggblade, Steven, and Gelson Tembo.** 2003. "Conservation Farming in Zambia." Environment and Production Technology Division Discussion Paper 108.
- Islam, Asad, Philip Ushchev, Yves Zenou, and Xin Zhang.** 2019. "The Value of Information in Technology Adoption: Theory and Evidence from Bangladesh." Unpublished.
- Jack, B. Kelsey.** 2011. "Market Inefficiencies and the Adoption of Agricultural Technologies in Developing Countries." Unpublished.
- Jackson, Matthew, and Evan C. Storms.** 2019. "Behavioral Communities and the Atomic Structure of Networks." Unpublished.

- Kim, David A., Allison R. Hwang, 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." *Lancet* 386 (9989): 145–53.
- Kremer, Michael, Edward Miguel, Sendhil Mullainathan, Clair Null, and Alix P. Zwane.** 2011. "Social Engineering: Evidence from a Suite of Take-up Experiments in Kenya." Unpublished.
- Lea, Nicholas, and Lucia Hammer, Lucia.** 2009. "Constraints to Growth in Malawi." World Bank Policy Research Working Paper WPS 5097.
- Magruder, Jeremy R.** 2010. "Intergenerational Networks, Unemployment, and Persistent Inequality in South Africa." *American Economic Journal: Applied Economics* 2 (1): 62–85.
- Miller, Grant, and A. Mushfiq Mobarak.** 2015. "Learning about New Technologies through Social Networks: Experimental Evidence on Nontraditional Stoves in Bangladesh." *Marketing Science* 34 (4): 480–99.
- Munshi, Kaivan.** 2004. "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution." *Journal of Development Economics* 73 (1): 185–213.
- Munshi, Kaivan.** 2007. "Information Networks in Dynamic Agrarian Economies." In *Handbook of Development Economics*, Vol. 4, edited by T. Paul Schultz and John Strauss, 3085–113. Amsterdam: Elsevier.
- Munshi, Kaivan.** 2008. "Social Learning and Development." In *The New Palgrave Dictionary of Economics*, edited by Lawrence E. Blume and Steven N. Durlauf. London: Palgrave Macmillan.
- Oster, Emily, and Rebecca Thornton.** 2012. "Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-Up." *Journal of the European Economic Association* 10 (6): 1263–93.
- Perla, Jesse, and Christopher Tonetti.** 2014. "Equilibrium Imitation and Growth." *Journal of Political Economy* 122 (1): 52–76.
- Qiao, Fangbin, Jikun Huang, and Xiaobing Wang.** 2015. "Fifteen Years of Bt Cotton in China: The Economic Impact and Its Dynamics." *World Development* 70: 177–85.
- Rogers, Everett M.** 1962. *Diffusion of Innovations*. New York: The Free Press.
- Rosenzweig, Mark R., and Christopher Udry.** 2020. "External Validity in a Stochastic World: Evidence from Low-Income Countries." *Review of Economic Studies* 87 (1): 343–81.
- Ryan, Bryce, and Neal C. Gross.** 1943. "The Diffusion of Hybrid Seed Corn in Two Iowa Communities." *Rural Sociology* 8 (1): 15–24.
- World Bank.** 2008. *World Development Report 2008: Agriculture for Development*. Washington, DC: World Bank.
- World Bank.** 2011. *World Development Indicators*. Washington, DC: World Bank.