

Leveraging Social Networks for Agricultural Development in Africa

Martha Ross

Thesis committee

Promotor

Prof. Dr E.H. Bulte
Professor of Development Economics
Wageningen University & Research

Co-Promotor

Dr M. Voors
Assistant professor, Development Economics Group
Wageningen University & Research

Other members

Prof. Dr J.H.H. Wesseler, Wageningen University & Research
Prof. Dr C. Leeuwis, Wageningen University & Research
Prof. Dr A. Kontoleon, University of Cambridge, UK
Dr B. D'Exelle, University of East Anglia, Norwich, UK

This research was conducted under the auspices of the
Wageningen School of Social Sciences (WASS)

Leveraging Social Networks for Agricultural Development in Africa

Martha Ross

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 13 June 2017
at 4 p.m. in the Aula.

Martha Ross

Leveraging Social Networks for Agricultural Development in Africa,
180 pages.

PhD thesis, Wageningen University, Wageningen the Netherlands (2017)
With references, with summary in English

ISBN: 978-94-6343-191-0 / DOI: 10.18174/413526

To my family, for all your love and support

TABLE OF CONTENTS

CHAPTER 1 Introduction	1
CHAPTER 2 Raising Farmer Yields in Eastern D.R. Congo.....	15
CHAPTER 3 Social Networks and Technology Diffusion.....	59
CHAPTER 4 Community Selection of Targets in Network Diffusion Interventions	99
CHAPTER 5 Social Networks and Social Preferences.....	127
CHAPTER 6 Conclusion.....	155
English Summary.....	167

CHAPTER 1

Introduction

1 Overview

Adam Smith noted in *The Theory of Moral Sentiments* (1759) that human exchange consists not just of goods and services, but also the trading of ideas, support and favors. Adam Smith's *invisible hand* was therefore driven not just by market transactions, but also by social exchanges. Such social exchanges were viewed by Smith as fundamental for the creation of solutions that benefited the community as a whole. When a new solution is introduced to a community, the extent of its impact rests in its diffusion through the community social network. With the digitalization of the modern world has come an explosion of high-resolution network data, allowing innovation diffusion and impacts to be quantitatively analyzed. Seizing upon these data-opportunities, research on social network diffusion has taken hold across the social science and business worlds (Pentland, 2014). However in 'offline' networks, data collection costs remain high and only recently has attention turned to exploring diffusion of innovations within developing societies. This thesis contributes to a growing literature that explores relationships between social networks and innovation diffusion within a developing country context. Given this context, the networks of interest within this thesis are the offline interpersonal relationships between community members. Diffusion channels for new innovation are therefore limited to word-of-mouth communication, observation, and personal experience.

While approaching these questions of how innovations diffuse through social networks, this thesis stays firmly rooted in the belief of idea exchange being a tool for creating positive community impacts. Using state of the art methods including field experiments, surveys, and simulations, this thesis asks how to better capitalize on how social networks can be leveraged for improving the effectiveness of practitioners in facilitating adoption of new innovations. While the chapters of this thesis provide interlinking insights on

this overarching question, each chapter is presented as a self-standing contribution to this emerging social network literature within development economics. This thesis begins by exploring current impacts of diffusion-based interventions on development outcomes before zooming in on critical design components within such interventions.

2 Learning and Development

Diffusion and social-learning based development initiatives have gained widespread popularity. Social-learning based programs offer the promise of reaching a widespread audience at lower cost to the implementing agent. While the specific design varies by intervention, many take the form of having professionals train a select number of community members who are then responsible for spreading learned information onwards through their social networks. Numerous studies have analyzed such programs and evaluated their effectiveness in catalyzing behavioral change in the fields of global health (Haider, Ashworth, Kabir, & Huttly, 2000; Kim et al., 2015; Kincaid, 2000; Oster & Thornton, 2012), microfinance (Banerjee et al., 2013; Bursztyn et al., 2014), insurance markets (Cai, 2012), and agriculture development (Beaman, BenYishay, William, & Mobarak, 2015; BenYishay & Mobarak, 2014; Conley & Udry, 2010; Kondylis, Mueller, & Zhu, 2014).

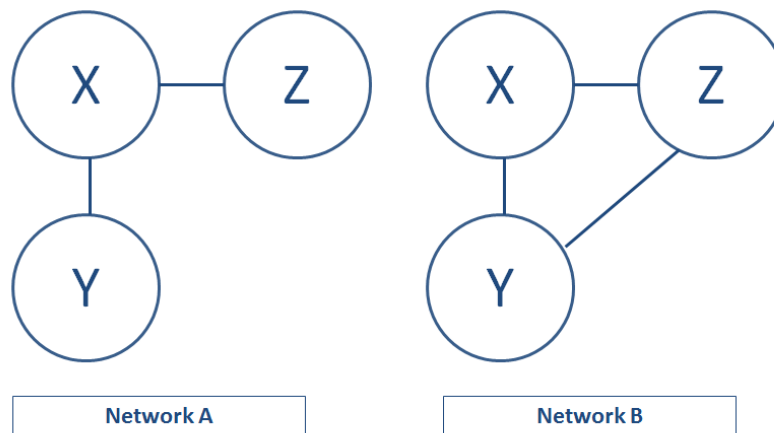
Agricultural development within Africa has been of particular focus as productivity-enhancing inputs are seen as crucial for raising farmer yields and incomes. This is largely due to the growing evidence of information gaps acting as a major constraint to the diffusion of improved inputs (Bandiera & Rasul, 2006; Munshi, 2003). The last decades have seen a rapid expansion of participatory agricultural extension programs which draw on this social learning frameworks. Many of these programs hinge on the argument that using community farmers as trainers improves knowledge transmission and learning because of the pre-existing relationships that external extension workers do not have (Simpson et al., 2015). These justifications highlight that practitioners recognize that information diffuses through social links and that social relationships effect influence. However only recently have these programs been approached and evaluated using more explicit network analysis methods (Beaman et al., 2015; Emerick, Janvry, Sadoulet, & Dar, 2016; Kim et al., 2015; Nourani, 2016).

In the context of agricultural innovations, motivations to adopt are largely tied to profit-maximization. However households in underdeveloped regions are largely independent from formal markets, making valuation of crop production and consumption difficult to quantify monetarily. As a result measuring returns on investment is often challenging for farmers and researchers alike. This information quality constraint contributes to slower adoption rates observed within sub-Saharan Africa (Feder and Rosenzweig, 2010). Similarly, agricultural technologies have largely heterogeneous returns to producers. Observation of neighbors' experiences with new innovations may be incorporated into adoption decision processes with lower confidence when prior adopters are not sufficiently comparable so as to provide a reliable benchmark of expected returns. In these situations, learning-by-doing may provide a more effective channel of learning. Direct experimentation would provide more accurate information on personal returns to adoption and be more influential in triggering adoption, assuming profitable returns are realizable.

3 Networks

The central focus of this thesis is the study of social networks as they relate to human decision-making and behavioral change. Before diving into why social networks are such an important component of understanding human decisions, the basic concepts of networks pertinent to this thesis are described. In analyzing social networks, the relationships that connect individuals are explored to explain variation in behavior and decisions (REF). In constructing network data first one must define the entity of interest ("node") and the connecting relationship(s) of interest ("links" or "edges"). By graphing networks, entire structures of nodes and links can be captured providing rich information on channels through which an individual may be shaped by social forces of interactions, information, and social sanctioning. Within this thesis nodes represent community households and links represent interpersonal relationships. Figure 1 provides a visual representation of a 3-node networks with two alternative link formations.

Figure 1 Simple Networks Demonstrating Connectedness



Network A of Figure 1 displays a network that is not completely connected while Network B is fully connected. In Network A, node X is linked to both Y and Z, but Y and Z are not linked with one another. In network terminology, this means that node X has a shortest path of 1 with both Y and Z while nodes Y and Z have a shortest path of 2 as two links must be travelled for them to reach one another (Y to X then X to Z). This is key to understanding information flows as it is the links between nodes that form the pathways by which information flows. In order for information from Y to reach Z, it must first go through X. However in Network B the network is completely connected, which means that all nodes are connected to all other nodes. The closer a network is to being completely connected the denser the network is considered to be. Density is an important measure as the structural level as information diffusion within denser networks is faster due to the numerous pathways by which the information can travel and reach all nodes (D. Centola, 2010; Choi, Kim, & Lee, 2010).

The level of connectedness at the node level is also important. Within Network A, node X lies on all pathways. That is, no pathways of links exist that do not include X. In terms of information flows, this means X will participate in all information flows. In Network B, nodes Y and Z can pass information without node X participating. As networks scale up, and the number of nodes and links increases, numerous approaches to calculating connectedness become possible. However the general rule of thumb is that the more connected an individual node is with other nodes, the more likely that they will be reached by network

information flows. Therefore by analyzing social network structures, and the position of individuals within those structures, much can be elucidated on an individual's participation within information flows and how that participation shapes behavior.

4 Diffusion of Innovations

When speaking of the diffusion of innovations, it is the rate at which innovations are adopted within a community or population that is often the metric of interest. However if the intent is to increase the adoption rate, one must disaggregate from the group down to the individual level and understand why some adopt and others do not. Adoption can be decomposed into two mutually-reinforcing stages: learning about the innovation through new information and the decision to adopt (or not) (Foster & Rosenzweig, 2010; Pentland, 2014). Learning about new innovations goes beyond an awareness of their existence, but constitutes the information set an individual has on the use of, returns to, and costs associated with the innovation. Individuals learn about new innovations through social learning or experiential learning. Social learning is when individuals source information through their interpersonal relationships. This may be through direct conversation, over-hearing conversations second hand, or observation of peers who have already adopted the innovation. Experiential learning, or learning-by-doing, is gained through personal experimentation with a new innovation. All sources of learning are incorporated into the information set on the new innovation and is used in determining whether adopting the innovation is expected to yield higher returns relative to the *status quo*. However the different information sources may be weighted and incorporated differently. Information received through social learning may be factored into the adoption decision both less accurately and more lightly as it comes with greater quality and applicability concerns, depending on its source, relative to first-hand experience (BenYishay & Mobarak, 2014; Hanna et al., 2014; Nourani, 2016). Importantly, non-adoption of an innovation does not preclude learning from having occurred, only that the new information did not result in adjusting expected returns so as to warrant adoption (Foster & Rosenzweig, 2010). Learning, and the spread of new information, is therefore a critical forum for understanding adoption of new innovations and their resulting impacts on societies.

Evidence suggests that information flows most strongly through the interpersonal relationships that make up the underlying social networks of communities (Foster & Rosenzweig, 2010; Hardy & Mccasland, 2015;

Pentland, 2014; Thomas W. Valente, 1996). Network structures therefore map out the pathways by which information flows. They are a major determinant of the learning and decisions processes that determine innovation adoption . An individual's position within their network structure will be highly determinant in the intensity with which they are exposed to new information and how that new information changes existing beliefs that drive adoption decisions. More frequent exposure to the new information increases the probability that individuals are aware of new innovations earlier and able to gain a more complete picture of the innovation's expected returns to adoption (Aral, Muchnik, & Sundararajan, 2009; D. Centola, 2010; D. M. Centola & Macy, 2007).

Not all exposure is equally impactful on the adoption decision, as certain network connections have a stronger social influence than others. Receiving new information from a socially influential network tie will have a greater likelihood of affecting behavioral change than information received from a non-influential tie. Social influence is not a static metric. It varies both within and across network connections. Individuals will be more susceptible to influence from those that they deem to be a domain expert, are similar in pertinent characteristic or behaviors, and with whom they feel a strong measure of trust (BenYishay & Mobarak, 2014; D. M. Centola & Macy, 2007; Pentland, 2014). The relationship between source and recipient in inciting adoption becomes more relevant for more complex innovations, or in instances where the knowledge gap between believed and real adoption returns is greater (Foster and Rosenzweig, 2010).

Both exposure and influence are closely correlated to the underlying social network structure and as a result network structures can be as influential in determining behavioral change as the content of information that is being spread. Network structures, and the flow of information, is constructed by interpersonal relationships and communication. In understanding diffusion dynamics and network structures thus hinges on channels of interpersonal communication within a given context.

Such channels of interpersonal communication vary widely between developed and under-developed regions of the world (Chuang & Schechter, 2000). In modern societies, the prevalence of mobile phones and online social media platforms have created communities that are at a massive-scale. This digitalization of network data has provided an unprecedented opportunity for network diffusion dynamics to be analyzed quantitatively with high-resolution data. This opportunity has been eagerly seized, with social diffusion

research growing rapidly and the emergence of social physics as an independent field (Pentland, 2015). A second result of digitalization is a fundamental restructuring of networks from local to global combined with an acceleration in the speed at which information is transmitted. Evidence suggests that diffusion dynamics in modern networks operate differently than in traditional network structures where information is less abundant and slower moving as ICT infrastructure remains largely absent (Choi et al., 2010; Dutta, Lanvin, & Paua, 2004). Social networks in these regions are formed and maintained through face-to-face interactions and word-of-mouth communication. Mapping social networks in the absence of digitalized communities is costly and time-intensive. Only recently have studies turned their attention to quantitatively analyzing the diffusion dynamics of these traditional 'offline' communities (Banerjee et al., 2013; Beaman et al., 2015; Kim et al., 2015).

Understanding adoption of new technologies within rural communities is best conducted at the community level. These community social networks can then be used to identify influential members who can better promote adoption of the technologies (Barahona and Pentland, 2007). By selecting these influential individuals as the injection points of new information into community social networks, faster-paced diffusion is better achieved. Once critical thresholds of adoption are achieved, mass adoption occurs and the new innovation transitions into common practice (Aral, Muchnik, & Sundararajan, 2013; D. M. Centola & Macy, 2007; Thomas W. Valente, 1996; Thomas W. Valente & Davis, 1999). Influential individuals have been found to cluster in the network, indicating that influential individuals are typically well connected to other influential network members. This clustering would make influential individuals prime targets as network entry points for increasing network adoption rates of new innovations and reaching thresholds of mass adoption (Valente and Davis, 1999; Aral and Walker, 2012). An individual's centrality within a network has been shown as a key characteristics for rapid diffusion and serves as a metric of identifying influential individuals. However numerous measures of centrality are available and identifying the appropriate measure by which entry points are selected must be tailored to the specificities of the technology in question.

Diffusion of simple technologies relies primarily on awareness and imitation. Use of simple technologies is easily achieved through observation and requires little additional training or focused knowledge for reaping returns. More complex technologies consist of greater knowledge intensity in order to capture returns to

adoption. Their use may require continued access to information and support on proper use in order for sufficient returns on investment to be successfully realized (Barahona and Pentland, 2007). Therefore, the more complex the technology, the more important the underlying network structure. Complex technologies require more frequent and repeated exposures as well as higher levels of confidence in the information being received in order to trigger adoption.

In determining the selection of individuals as network entry-points, the knowledge-intensity of a technology must be factored in along with the desired network positioning of the targeted entry-points. Centrality is a key metric in understanding the power and influence that a given individual has within their community social network. Numerous measures of centrality exist and determining the proper one is dependent on the type of information being spread and the objectives of the program (Borgatti, 2005; T. W. Valente, 2012). Recursive measures, such as eigenvector centrality, capture how well connected individuals are to other highly connected individuals, positioning it as a measure of influence. Chapter 3 of this dissertation explores whether diffusion of new technologies varies substantially when network injection points are selected based on eigenvector centrality.

5 Network Diffusion and Development

Development practitioners relying on social learning mechanisms, as seen within agricultural extension programs, highlights an implicit recognition of information diffusion through interpersonal relationships. Harnessing diffusion dynamics to strengthen social learning processes could generate positive impacts on development program effectiveness. Unfortunately, the data demands for network-based intervention implementation may be prohibitively high for development practitioners. Evidence suggests that network members have implicit, but incomplete, information on the network structures within which they are embedded (Banerjee et al., 2014). Chapter 4 explores whether a cost-effective means of harnessing network-dynamics can be found by aggregating network members incomplete information sets through a community-wide nomination process of contact-points. Finally, social learning programs require network contact-points be sufficiently pro-social so as to incur some individual costs in sharing information onwards, for the benefit of the wider community. Chapter 5 explores relationships between pro-social preferences with network centrality.

6 Objectives

Social networks provide a major channel through which information diffuse through populations. While information on new innovations is a critical component of the decision to adopt, the diffusion process of innovations is more complicated than just the spread of information. Technology adoption requires that an individual believes the returns of the new technology outweigh returns of existing technologies. This dissertation explores dynamics of interpersonal diffusion through social networks within the context of agricultural development of small holder crop producers in eastern Democratic Republic of Congo. While the independent chapters of the dissertation are academic in their methodologies and the specificity of their research questions, the aim of the thesis as a whole is to apply the insights from these chapters to development initiative design. The overarching objective of this dissertation is to identify opportunities within the design and implementation of development initiatives for improved harnessing of social network diffusion dynamics.

Chapter 2 Can social learning and subsidized personal experimentation increase farmer crop output levels?

Chapter 3 Does selection of network entry points using eigenvector central individuals improve technology diffusion?

Chapter 4 Can community nomination of entry points provide a cost-effective alternative to network-based entry point selection?

Chapter 5 Is an individual's position with their network related to demonstrated pro-social preferences?

Chapter 2 provides a snapshot on an existing diffusion-based intervention design to explore its effectiveness in accomplishing objectives of improving farmer welfare via higher crop output. This intervention uses traditional extension program selection methods of imposing socioeconomic criteria and selecting network entry points from eligible candidates. Chapters 3 and 4 dive into this question of network entry selection and effective diffusion by exploring alternative selection methods. Chapter 3 focuses on selection based solely on network parameters of eigenvector centrality while Chapter 4 assesses less data-costly alternatives of harnessing community knowledge on network structures through an open nomination process. Chapter 5 explores the relationship of social preferences and network centrality. In better understanding the

relationship between these traits, a clearer picture emerges on the links between exposure and social influence for better selection of network contact points. Taken in combination, the chapters provide an overview of available network dynamics that can be manipulated within intervention designs for generating stronger diffusion program impacts.

7 Methodology

7.1 Experimental Methods

Development economics has a lengthy history of collecting original data and testing specific economic theories through quantitative analysis. The size and quality of datasets within developing country continue to grow as technologies enable lower-cost data collection, allowing for richer datasets to be used within such empirical tests. However observational data makes identification of causal relationships empirically challenging. In efforts to better identify causal relationship, development economists have pulled on methodological approaches applied in other fields of research, most notably the fields of medicine, psychology, labor, and industrial organization (Duflo, 2006). This thesis draws primarily on three of these experimental methodologies, randomized controlled trials, field experiments and lab-in-field experiments.

Impact evaluations based on randomized controlled trials offers an opportunity for measuring causal, rather than just correlational, relationships between development interventions and behavioral change. Impact evaluations are conducted in natural environments with participants that may be unaware of their decisions being studied (Barrett & Carter, 2010; Gneezy & Imas, 2017). Impact evaluations are most effective in testing the effectiveness of solutions to specific practical problems within a given setting. The evaluation therefore tests prior theories on the most effective means of eliciting particular behavioral changes for overcoming the problem in question (Banerjee & Duflo, 2009; Barrett & Carter, 2010; Duflo, 2006). Chapter 2 of this thesis employs a quasi-experimental design in testing the impact of social learning and reduced-cost access to experimentation on increasing technology adoption. These interventions are specifically designed to the context and obstacles that farmers in eastern Democratic Republic of Congo are facing. While impact evaluations offer causal identifications of intervention impacts on development outcomes, they speak less to the underlying mechanisms that explain behavioral change.

Lab-in-field experiments are a marriage of the two approaches, offering more controlled measurement of behavioral responses, moving standardized, and validated, laboratory designs to more naturalistic environments (Gneezy & Imas, 2017). Lab-in-field experiments are specifically used to measure individual preferences and behavioral responses under highly controlled conditions. As a result lab-in-field experiments allow direct comparison of results across geographies and populations (Harrison & List, 2004). However the extent to which lab-measured behaviors correspond to ‘real world’ behavior remains contested (Andersen, Harrison, Lau, & Rutström, 2010; Berkowitz & Donnerstein, 1982; Levitt & List, 2007; Roe & Just, 2009). Chapter 5 of this thesis utilizes lab-in-field experimental methodologies in eliciting the social preferences of community members. Specifically, we use variations of the standard investment game and public goods game.

Framed field experiments provide a marriage of the two above experimental methodologies. They offer a more controlled environment to test behavioral responses to solutions addressing specific development problems while providing insights on underlying mechanisms. Chapter 3 of this thesis uses a framed-field experiment in order to measure the difference in diffusion patterns resulting from two different types of network entry-points: most eigenvector central versus least eigenvector central.

7.2 Network Analysis

Social institutions lie at the heart of each chapter within this thesis. A the most important institution measured within this thesis is that of the underlying network structures of the sample communities. Unlike network analysis that is conducted on digital networks, existing data on community networks in remote regions of the world is extremely sparse. All network data included within this thesis is original and collected through intensive surveying. In mapping networks through surveys, critical concerns that were factored into survey designs were the structure of relationship questions, which relationships to map, and measurement error resulting from memory bias.

The network mapping strategy used throughout this thesis pulls strongly from the approach used in Banerjee et al. (2013). Deviations from their approach include fewer relationship types being measured and the format in which responses were collected. Unlike Banerjee et al. (2013), no upper limit was imposed on the number of names listed per relationship type. Unlimited self-reporting of network ties intrinsically

introduces data-quality concerns of memory bias. As a safe-guard against this, a full census of households was read off once for each relationship of interest. Respondents were asked to respond “yes” or “no” with each name, indicating whether the relationship was present or not between them. The trade-off with this approach is survey-fatigue, as repeatedly reading off long-lists of names is time-intensive and fails to hold the interest of the respondent. To minimize risks of survey-fatigue, network mapping was limited to three dimensions found to be most important for the agricultural context. While incomparable to the rich network datasets generated by digitalized networks, these network maps provide a novel glimpse into the social lives of rural communities. The dataset used in this thesis is one of only a few that measures nearly complete community networks within such an underdeveloped region of the world.

8 Outline

The thesis is structured as follows. Chapter 2 uses a quasi-experimental design in evaluation two agricultural interventions on raising farmer crop output in eastern Democratic Republic of Congo. The first treatment arm relies only on a social learning design while the second combines the social learning program with an additional layer of reduced-cost experimentation in the form of an offer for subsidized packages of improved inputs. Chapter 3 studies diffusion processes of in-organic fertilizer and associated information on the use and returns to in-organic fertilizer application. Specifically it explores whether diffusion can be improved by selecting entry points into community networks based on eigenvector centrality. Chapter 4 tests whether community-selection of injection points offers a comparably alternative for improving diffusion without necessitating the same prohibitive data-demands of network-based selection. Chapter 5 studies whether pro-social values of trust, trustworthiness, and cooperation are associated with network position. Incentivized lab-in-field experiments are used to elicit these social preferences of a sub-sample of community members. These community members are drawn based on eigenvector centrality scores within their respective networks. Chapter 6 concludes by discussing the main findings of this thesis and the cross-cutting lessons gleaned from the individual chapters. In addition, the limitations of the thesis are reviewed and recommendations for future research are suggested.

References

- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2010). Preference heterogeneity in experiments: Comparing the field and laboratory. *Journal of Economic Behavior and Organization*, 73(2), 209–224. <https://doi.org/10.1016/j.jebo.2009.09.006>
- Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America*, 106(51), 21544–21549. <https://doi.org/10.1073/pnas.0908800106>
- Aral, S., Muchnik, L. E. V., & Sundararajan, A. (2013). Engineering social contagions: Optimal network seeding in the presence of homophily. *Network Science*, 1(2), 125–153. <https://doi.org/10.1017/nws.2013.6>
- Bandiera, O., & Rasul, I. (2006). SOCIAL NETWORKS AND TECHNOLOGY ADOPTION. *The Economic*, 116(514), 869–902.
- Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013). The diffusion of microfinance. *Science*, 341, 363–370. <https://doi.org/10.1126/science.1236498>
- Banerjee, A. V., & Duflo, E. (2009). The Experimental Approach to Development Economics. *Annual Reviews of Economics*, 1, 151–178. <https://doi.org/10.1146/annurev.economics.050708.143235>
- Barrett, C. B., & Carter, M. R. (2010). The power and pitfalls of experiments in development economics: Some non-random reflections. *Applied Economic Perspectives and Policy*, 32(4), 515–548. <https://doi.org/10.1093/aep/p000>
- Beaman, L., Benyishay, A., William, C., & Mobarak, A. M. (2015). *Can Network Theory based Targeting Increase Technology Adoption?*
- BenYishay, A., & Mobarak, A. M. (2014). Social Learning and Communication. *NBER Working Paper*, (20139). <https://doi.org/10.3386/w20139>
- Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist*, 37(3).
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27, 55–71. <https://doi.org/10.1016/j.socnet.2004.11.008>
- Bursztyn, L., Ederer, F., Ferman, B., & Yuchtman, N. (2014). Understanding Mechanisms Underlying Peer Effects: Evidence from a Field Experiment on Financial Decisions. *Econometrica*, 82(4), 1273–1301.
- Cai, J. (2012). Social Networks and the Decision to Insure : Evidence from Randomized Experiments in China *.
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329, 1194–1197. <https://doi.org/10.1126/science.1185231>
- Centola, D. M., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3), 702–734. <https://doi.org/10.1086/521848>
- Choi, H., Kim, S.-H., & Lee, J. (2010). Role of network structure and network effects in diffusion of innovations. *Industrial Marketing Management*, 39(1), 170–177. <https://doi.org/10.1016/j.indmarman.2008.08.006>
- Chuang, Y., & Schechter, L. (2000). Social Networks in Developing Countries. *Annual Review of Resource Economics*, (7). <https://doi.org/10.1146/annurev-resource-100814-125123>
- Conley, B. T. G., & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *The American Economic Review*, 100(1), 35–69.
- Duflo, E. (2006). Field Experiments in Development Economics 1. *Econometric Society Monographs*, 42(12), 1–30. Retrieved from <http://qed.econ.queensu.ca/pub/students/phds/khana/econ239/readings/duflo.pdf>
- Dutta, S., Lanvin, B., & Paua, F. (2004). *The Global Information Technology Report 2003-2004 Towards an Equitable Information Society*.
- Emerick, K., Janvry, A. De, Sadoulet, E., & Dar, M. H. (2016). *Identifying early adopters , enhancing learning , and the diffusion of agricultural technology*.
- Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of Technology Adoption. *Annual Review of Economics*, 2. <https://doi.org/10.1146/annurev.economics.102308.124433>
- Gneezy, U., & Imas, A. (2017). Lab in the Field: Measuring Preferences in the Wild. *Handbook of Field Experiments*, Forthcomin, 1–40. <https://doi.org/10.1017/CBO9781107415324.004>
- Haider, R., Ashworth, A., Kabir, I., & Huttly, S. R. (2000). Effect of Community-Based Peer Counsellors on Exclusive Breastfeeding Practices in Dhaka, Bangladesh: A Randomised Controlled Trial. *Lancet*, 356(9242), 1643–1647.
- Hanna, R., Schwartzstein, J., Benjamin, D., Gentzkow, M., Halac, M., Newman, A., & Ogden, T. (2014). Learning Through Noticing: Theory and Experimental Evidence in Farming.
- Hardy, M., & McCasland, J. (2015). *It Takes Two : Experimental Evidence on the Determinants of Technology Diffusion*.
- Harrison, G. W., & List, J. a. (2004). Field Experiments. *Journal of Economic Literature*, 42(4), 1009–1055.
- Kim, D. A., Hwong, A. R., Stafford, D., Hughes, D. A., O'Malley, A. J., Fowler, J. H., & Christakis, N. A. (2015). Social network targeting to maximise population behaviour change: A cluster randomised controlled trial. *The Lancet*, 386, 145–153. [https://doi.org/10.1016/S0140-6736\(15\)60095-2](https://doi.org/10.1016/S0140-6736(15)60095-2)
- Kincaid, D. L. (2000). Social Networks, ideation, and contraceptive behavior in Bangladesh: A longitudinal analysis. *Social Science & Medicine*, 50(2).

- Kondylis, F., Mueller, V., & Zhu, S. J. (2014). *Seeing is believing? Evidence from an extension network experiment* (Policy Research Working Paper No. 7000).
- Levitt, S. D., & List, J. A. (2007). About the Real World ? *Journal of Economic Perspectives*, 21(2), 153–174.
- Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *The Quarterly Journal of Economics*, (May), 549–599. <https://doi.org/10.1162/003355303321675455>
- Nourani, V. (2016). *Social Network Effects of Technology Adoption : Investing with Family , Learning from Friends & Reacting to Acquaintances*.
- Oster, E., & Thornton, R. (2012). Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-Up. *Journal of the European Economic Association*, 10(6), 1263–1293. <https://doi.org/10.1111/j.1542-4774.2012.01090.x>
- Pentland, A. (2014). *Social Physics: How good ideas spread-the lessons from a new science*. Penguin.
- Roe, B., & Just, D. (2009). Internal and External Validity in Economics Research: Tradeoffs between Experiments, Field Experiments, Natural Experiments, and Field Data. *American Journal of Agricultural Economics*, 91(5), 1266–1271.
- Simpson, B. M., Franzel, S., Degrande, A., Kundhlande, G., & Tsafack, S. (2015). Farmer-to-Farmer Extension: Issues in Planning and Implementation. *MEAS Technical Note*, (May).
- Smith, A. (1759). *The Theory of Moral Sentiments*.
- Valente, T. W. (1996). Network models of the diffusion of innovations. *Computational & Mathematical Organization Theory*, 2(2), 163–164.
- Valente, T. W. (2012). Network Interventions. *Science*, 337(6090), 49–53. <https://doi.org/10.1126/science.1217330>
- Valente, T. W., & Davis, R. L. (1999). Accelerating the Diffusion of Innovations using Opinion Leaders. *The ANNALS of the American Academy of Political and Social Science*. <https://doi.org/10.1177/0002716299566001005>

CHAPTER 2

Raising Farmer Yields in Eastern D.R. Congo

Abstract

Yields in Africa remain persistently and substantially below potential levels of output. More intensive use of improved agricultural inputs has been put forward as a strategy for raising farm productivity and improving livelihoods of smallholder farmers. However adoption of improved inputs within the region has been slow. Considerable evidence points to the role of information gaps in the low adoption of productivity-enhancing inputs in perpetuating these low yields. This study analyses two policy tools in targeting these information gaps. The first is through social learning as part of a farmer extension program. The second combines social learning with experiential learning, reducing the cost to personal experimentation with subsidized improved input packages. Our results indicate that farmers who are exposed to both social learning and learning-by-doing more significantly impacts farmer productivity relative to those receiving no intervention and those exposed only to social learning.

Publication Status: Ross, M. 2017. “Raising Farmer Yields in Eastern D.R. Congo: social learning and reduced-cost experiential learning on farmer productivity.” Working Paper.

1 Introduction

Amidst increasing concerns of food security, high rural-to-urban migration, and persistent poverty within sub-Saharan Africa, policy makers have become increasingly focused on improving currently low output levels of smallholder farmers (World Bank, 2007; FAO, 2009; Godfray et al., 2010; Irz, Lin, Thirtle, & Wiggins, 2001; Rosegrant & Cline, 2003). While farmers in developed regions of the world achieve an average of 80% of potential crop yields, farmers in sub-Saharan Africa experience rates closer to 20% (Titttonell & Giller, 2013). This large gap between realized and potential yield levels is thought to best be closed through more intensive use of productivity-enhancing inputs such as improved seeds, in-organic fertilizer, and improved pest and disease management systems, however adoption rates of these technologies remain very low. Considerable evidence points to the role of information gaps in the low adoption of productivity-enhancing inputs within the region (Bandiera & Rasul, 2006; Conley & Udry, 2010; Kondylis et al., 2014). Additional constraints on output and agricultural technology adoption include missing credit, input, and output markets and poor soil fertility management (USDA, 2013; Fafchamps et al., 2005; World Bank, 2008; Denning et al., 2009; Ehui & Pender, 2005). These constraints both independently and in combination obstruct agricultural development within the area and perpetuate low crop yields.

Current barriers to agricultural development stem directly from constraints, both resource and knowledge, faced by farmers and indirectly from larger institutional and infrastructural contexts within which farmers make their production decisions. Direct constraints would consist of farmer knowledge gaps on improved production processes and management systems and insufficient financial resources, both in terms of cash endowments and access to credit, for adopting productivity-enhancing inputs. These direct constraints are both more targetable and more easily manipulated in the short term, but face limitations in their scope of impact as persistent shifts in production require that higher productivity yields economic returns. When institutional indirect effects prevent farmers from reaping profits from productivity-enhancing technologies, farmers have little incentive to adopt the technologies and shift their production processes. Institutional constraints can take the form of high market transaction costs, an absence of credit or insurance markets, missing value chain linkages between local and export markets, insecure property rights, and unreliable governing institutions. These indirect constraints are larger and slower-moving barriers that require shifts

in the institutional framework of a region, making them difficult to target with external development programs and initiatives.

Ultimately, the combination and type of constraints a farmer faces will determine what level of production is efficient as well as how effective interventions that address different constraints will be in shifting the farmer to a new, higher, efficient output level. This study explores two policy instruments of choice for development practitioners that target constraints of information to improved inputs in inducing technology adoption and raising farmer output levels. The first is an information-transfer approach while the second seeks to lower the costs associated with learning-by-doing through experimentation. This study uses a quasi-experimental design in evaluating a purely information-transfer approach against a combined information-transfer with lowered transaction costs to experimentation on raising farmer crop yields. Section 2 presents the implicit theory that underlies information-transfer and transaction-cost targeting agricultural development programs as well as existing empirical studies that evaluate the effectiveness of farmer extension and subsidization projects on outcomes of technology adoption and farmer productivity.

The intervention consists of two treatment arms. The first treatment arm provides farmers with agricultural extension services covering soil fertility management, legume production (in particular common beans and soybeans), in-organic fertilizer use, and improved farm management methods. The second treatment combines this same agricultural extension service with an additional offer for subsidized input packages. The first treatment thus targets knowledge gaps independently while the second targets both knowledge gaps and transaction cost to experimentation simultaneously. Both treatment arms are compared to a control group that receives neither intervention.

Our study proposes to contribute to existing literature in three ways. First, those farm and household characteristics that contribute to greater farmer crop output in a remote and under-developed region of sub-Saharan Africa are explored. These characteristics can provide insights on channels through which agriculture development programs can be effectively targeted. Second, a quasi-experimental methodology is used to analyze synergistic benefits of addressing both information and financial constraints simultaneously through a dual-treatment structure. Farmers are randomly selected to receive either farmer-led agricultural extension alone or agricultural extension combined with a subsidized input-package offer.

Using a difference-in-difference model I am able to isolate the effect of the extension program independently and compare results to the impact of the combined extension and subsidy program on farmer technical efficiency and total output value. This design effectively evaluates a purely social learning design versus a combined social learning with reduced cost experimentation in identifying a more effective means of addressing information constraints to agricultural technology adoption. Third, I explore which production decisions may have been most influenced by the two interventions, and are therefore more malleable facets of smallholder production which agricultural policy can target within intensification programs.

Results indicate that addressing a combination of social learning with reduced-cost experimentation simultaneously generates larger benefits in terms of raising overall farm output compared to providing only the social learning-based extension program. Model estimates show a significant impact for only the combined extension-plus-subsidy treatment arm on total value of farm output. In exploring channels through which the intervention may have effected crop yields, households in this treatment arm are found to be more likely to change their intercropping patterns, to reduce their overall input use between time periods, and to be more likely to adopt in-organic fertilizer on their farm. The reduction in total input use suggests farmers are realizing synergistic benefits of adopting improved inputs and management processes that allow greater output for small input bundles.

The remainder of the paper is structured as follows. Section 2 positions this study within the existing literature on agricultural extension, social learning, and technology adoption. Section 3 provides a detailed overview of the study design, specifically the two interventions evaluated within this study. Section 4 outlines the empirical model used in evaluating intention to treat effects of the two interventions. Additionally, Section 4 presents empirical models used to explore potential mechanisms driving estimated treatment effects. Specifically, I analyze the impact of the interventions on production decisions of intercropping behavior, total inputs used on farm, yield levels for common beans, and uptake in in-organic fertilizer. Section 5 provides a summary of the data collected while Section 6 catalogues estimated results. Section 7 concludes.

2 Literature Review

2.1 Information within Technology Adoption Decisions

Despite productivity-enhancing technology adoption being identified as being a critical channel for agricultural development within sub-Saharan Africa, adoption rates in the region remain low (Kondylis et al., 2014; World Bank, 2008). Knowledge gaps have been identified as a major contributor to the low adoption rates, prompting initiatives to better diffuse information around the returns on investment for and proper use of improved inputs (Bandiera & Rasul, 2006; Conley & Udry, 2010; Kondylis et al., 2014). Agricultural technology adoption is largely prompted by expected returns on investment for the new technology being greater than returns on investment of existing technologies. A seemingly simple relationship, identifying and quantifying returns in practice is complex and knowledge factors into the adoption decisions in several forms (Foster & Rosenzweig, 2010).

Knowledge acts as a major determinant of the shape of a household's production function as it drives the perception a farmer has of the expected output levels for a given allocation of inputs across produced outputs. Knowledge may shape the production function through multiple forms. It may be in the skill with which a factor is applied within the production process, in the perception of output variability given a set of inputs, or in information on the prices and transaction costs for markets which a farmer uses in determining their market participation levels and the proportion of output to consume versus sell. Limited information can thus create a divergence between the distributions of actual versus perceived net benefits in several key areas of production decision making, causing uptake levels to be held artificially low as farmers misestimate expected overall returns to adoption (Feder & Slade, 1984).

Sources of knowledge can come through direct knowledge transfer (e.g. formal trainings), observation of prior adopters, or personal experimentation. The appropriate combination of sources for most effectively aligning farmer perceptions of returns is highly variable depending on the technology in question. Direct knowledge transfers are more effective for technologies that are more knowledge-intensive and whose returns on investment are more sensitive to quality of application. While formal trainings can provide general information on application techniques and average returns to adoption, they are not necessarily able to provide farmer-specific information on how to best achieve these productivity gains given the farmer's input

bundles, crops of production, and farm characteristics (Foster & Rosenzweig, 2010). Learning through observation of prior adopters provides a sample of more heterogeneous examples of returns to adoption without necessitating costs of investment to the observer. However learning is still incomplete as perfect information on complementary aspects of production and the integration of the new technology (e.g. interactions between the technology and soil quality) is impossible through observation alone (Maertens & Barrett, 2012). Therefore for technologies that have highly heterogeneous returns, learning-by-doing provides more complete information both on expected returns and optimal allocation and application of inputs with the new technology for achieving highest returns given the farmer's production conditions. However for technologies that must be purchased through markets, experiential learning requires some investment and introduces potential financial barriers.

As knowledge endowments evolve through experience, observation, or formal training, expectations around output variability is reduced and, under conditions where real positive returns to investment exist for the farmer, the production function under the new technologies shifts from low to high perceived net returns. This shift in expected returns increases the likelihood of both whether and how intensely a new innovation is adopted (Hiebert, 1974).

2.2 Determinants of Technology Adoption

While more accurate information on the returns and proper use of improved inputs is a critical component of decisions to adopt new technologies, it purely shapes the beliefs of farmers on the profitability of adoption. Action depends on farmers beliefs changing sufficiently that a switching threshold, at which the net returns to the new innovation are believed to be greater than the new returns of the existing technology. For smallholder farmers who both consume and market their output and produce primarily with household labor and self-produced inputs, identifying the net returns of new technologies is not straight forward. This section outlines the general framework within which agricultural households make their crop production decisions.

Agricultural households are characterized by the non-separability of their consumption and production decisions in the allocation of their fixed set of input resources (Benjamin, 1992; Ellis, 1993; Taylor & Adelman, 2003). This interdependency between production and consumption makes it necessary that

farmers optimize not solely based on production functions but instead optimize both consumption and production simultaneously under a single utility function (Benjamin, 1992). Households then maximize utility with simultaneous production and consumption decisions subject to constraints of available cash, productive asset endowments, family time and preferences between labor and leisure, production technologies, and prices of inputs, outputs, and non-produced consumption goods (Feder & Umali, 1993; Taylor & Adelman, 2003). In constructing agricultural household production functions the relationships of factor-output, factor-factor, and output-output are critical in shaping farmer production expectations over which they maximize utility (Ellis, 1993). In the simplest models, farmers decide to adopt a new technology by estimating the profitability of adoption based on three key production relationships: factor-output, factor-factor, and output-output. Optimal use, and maximal profits, rely on each of these relationships being efficiently balanced; a process typically achieved with information gained through experimentation, observation of prior adopters, and formal knowledge transfer channels.

The factor-output relationship captures the transformation of a given input into generated output. Innovations that increase the farmer's yield for a given level of inputs are said to be productivity enhancing. Factor-factor relationships captures the combinations of inputs that generate a given level of output. Factors are often interrelated and can be conceptualized as packages of complements. New technologies cannot therefore be evaluated purely in their independent relationship to outputs, but their relationships with other existing factors of land, labor, and farm size must also be considered (Ellis, 1993). The linkages between a technology and other factors of production will determine the extent to which farmers are able to realize synergistic benefits and the how the shape of the production function changes with the introduction of a new technology. Things become more complicated as multiple new technologies become available that have positive synergistic relationships with one another. With multiple technologies the probability of both technologies being adopted simultaneously increases with stronger synergistic benefits, higher divisibility of the two technologies, and lower transaction costs in acquiring the technologies and selling extra generated output (Feder & Umali, 1993).

Finally, output-output relationships are relevant as they determine how to allocate a fixed bundle of input resources across output activities. This relationship becomes most pertinent when innovative technologies

are output specific or when the technology is a new output activity. Each of these provides economic incentives for a re-allocation of the household's fixed set of inputs while simultaneously accounting for the factor-output and factor-factor relationships that are relevant for the optimal allocation of resources to generate expected outputs that will maximize household utility.

Developing countries face infrastructure and exacerbated production uncertainty that restricts agricultural households investments in new technologies as adjustments in product (food) output and factor (labor and land) utilization is constrained (Janvry, Fafchamps, & Sadoulet, 1991). As a result, practicalities of missing markets and household risk must be incorporated into models of agricultural decision making as obstacles in capturing potential returns to new technologies are present. Numerous models were developed to more precisely reflect real world realities that prevent farmers from realizing potential returns on investments in improved inputs and therefore contribute to observed low adoption rates of these inputs (Feder & Umali, 1993).

A market is defined as “missing” if the cost of market-participation is sufficiently high that the optimal household strategy is one of self-sufficiency. Transaction costs that decrease profitability of produced goods and raise purchase prices of consumed goods create price bands in which a household is best off withdrawing from formal markets and becoming self-sufficient (Key, Sadoulet, & Janvry, 2000). As transaction costs are not uniform, households have varying levels of integration into product, factor, and labor markets. The more integrated into the market, the more responsive a household is to productivity-enhancing technologies as the differences in market prices and a household's shadow price is more easily reduced to the point households can capture returns on their excess output (Key et al., 2000). As a result, any price-targeting policies within these markets will have ambiguous effects on technology adoption decisions depending on whether transaction costs outweigh expected returns to adoption (Taylor & Adelman, 2003). High transaction costs of acquiring new innovations limit self-experimentation, a more direct information source on the suitability of new technologies for specific farm conditions and potential productivity returns.

In addition to facing imperfect market integration, agricultural households in developing countries also face high uncertainty. Crop production as mitigating strategies are limited and market prices are highly responsive

to variations in supply, an interlinkage that can create multiplicative effects which households have to account for within their production decisions. In the face of sunk costs to production and high uncertainty, farmers prioritize flexibility over scale and lower risk over higher potential maximum output levels in making their adoption decisions (Dixit & Pindyck, 1994; Hiebert, 1974). Risk aversion has been found to be negatively correlated to a farmer's wealth. As wealthier farmers have a larger buffer against failure, they're more willing to engage in more productive, but higher risk, technologies (Just & Zilberman, 1983). This relationship between risk-aversion and wealth can result in risk-induced poverty traps as marginal producers are restricted to less profitable technologies, effectively trapping them in poverty (Dercon & Christiaensen, 2008; Fermont, van Asten, Titttonell, van Wijk, & Giller, 2009; Titttonell & Giller, 2013). Tendencies of farmers to prioritize low risk investments highlights the need for farmers to have access to accurate information on the potential returns to investment. Farmer uncertainty is exacerbated by farmers having highly variable expectations around technology performance and effects on productivity.

2.3 Agricultural Extension Interventions

While increasing farmer productivity is the decisive goal of policy makers, division remains over the best channel through which to develop agriculture and increase farmer output levels (Christiaensen, Pan, & Wang, 2010)). Many agricultural development policies push for adoption of yield-increasing inputs such as improved seeds, in-organic fertilizer, and greater soil, pest, and disease management processes. Adoption of high-yielding genotypes has been widespread through sub-Saharan Africa but adoption of in-organic fertilizer and new farm management processes have failed to take off (Foltz, Aldana, & Laris, 2012; Morris, Kelly, Kopicki, & Byerlee, 2007; Nisrane et al., 2011; Zingore, 2011). These low adoption rates are frequently attributed to two primary sources: knowledge gaps and market constraints. Agricultural extension has become a popular tool for targeting knowledge gaps in efforts to increase adoption rates of improved inputs.

Our study is designed to assess the impact of an agricultural development project that uses farmer extension and subsidized input packages in addressing knowledge transfer and cost-to-experimentation constraints. Extension programs often involve a participatory approach that provide farmers with information on-, practical training with-, and opportunities to observe experimental plots with improved inputs and management practices. Isolating impact effects of extension programs on farmer productivity is complicated

by a complex chain of causation that includes knowledge, adoption of better practices, and investment in improved inputs. Despite these causal-chain challenges, rigorous evaluation of extension programs is necessary to determine whether returns to beneficiaries sufficiently justify the high program implementation costs.

To date, agriculture extension evaluations yield mix results (for a review, see Anderson & Feder, 2007). Many evaluations are riddled with issues of data quality, measurement error, and methodological weaknesses that undermine causality. However, several studies have emphasized the importance of program design in targeting local production constraints and farmer needs. One area of particular focus is the extent to which agricultural extension services are currently pro-poor. Several studies suggest that program benefits are being disproportionately accrued by wealthier farmers (Agyei-Holmes, Ayerakwa, Osei, & Osei-Akoto, 2011; Cerdán-Infantes, 2008; Cunguara & Moder, 2011).

The second constraint to agricultural development this study specifically targets is cost to self-experimentation with new technologies. Weak or missing financial markets result in poor access to credit and insufficient funds to purchase improved inputs while access to input markets are often associated with distortionary transportation costs, driving up input prices and undermining the economic viability of improved-input use for remote farmers. Growing new or failing markets into functional institutions is beyond the scope of development practitioners so market constraints are typically targeted through subsidization. Subsidy schemes vary in their precise structure and scope, but all are intended to lower existing price barriers sufficiently to open markets up to farmers with limited resources.

Subsidization of agricultural inputs is still widely debated in its effectiveness as a policy tool. Proponents of subsidization argue that subsidies encourage adoption and/or optimal use of improved inputs by addressing constraints arising from risk aversion and market imperfections (Crawford & Kelly, 2006; Druilhe & Barreiro-hurlé, 2012). Critiques argue that subsidies increase demand for inputs, driving up market prices and making sustainability of the programs costly for governments. Critiques also highlight that subsidies are susceptible to political corruption and could crowd out commercial distribution channels (Crawford & Kelly, 2006; Druilhe & Barreiro-hurlé, 2012).

Studies evaluating the returns to subsidization schemes contain mixed results on the success of input subsidization in kick-starting agriculture development. Similar to the extension program evaluation studies, several subsidy-impact studies emphasize the necessity for smart designs. The importance of training on proper input use, timing of subsidy offers, and targeting of beneficiaries to avoid disproportionate capture by wealthy farmers have all been highlighted as key considerations for maximizing effectiveness (Carter, Laajaj, & Yang, 2014; Duflo, Kremer, & Robinson, 2011; Marenja & Barrett, 2009). However no clear consensus has emerged on whether subsidization positively impacts agricultural intensification efforts.

This study focuses on a one-off subsidization offer for the purpose of allowing farmers to experiment with a selection of improved inputs on their own fields for a single growing season. It is likely that interaction effects between knowledge transfer through social learning and personal experimentation exist. This study aims to better understand such interaction effects and whether the additional costs of providing subsidized inputs for experimentation is justified by greater gains to farmer output and productivity.

3 Research design

3.1 South Kivu Province

This study is set in South Kivu province of eastern DRC, a region that faces entrenched poverty and persistent armed conflict insecurity. Central government institutions are only weakly present with localized governance institutions providing the main structure of political daily life. Infrastructure in the region is under-developed, resulting in poor access to markets and high transportation costs in moving goods. The local population relies primarily on subsistence agriculture with limited off-farm wage Labor opportunities. Agricultural production takes place over two growing season, a short season from February through June and a longer season from July through December . Despite two growing periods and supplementation of some off-farm wage Labor, current estimates are that greater than 60% of regional households are food insecure. Food insecurity challenges are compounded by high regional population growth rates and limited supply of still un-cultivated arable land. The combination of these two forces creates persistent excess demand for land and rapid soil degradation (CIALCA, 2007).

Topographically and agri-ecologically, the region is quite varied across the large province. Much of the land is characterized as tropical highlands with average altitude in the area of 1,500 meters above sea level (CGIAR, 2014; Maass, Musale, Chiuri, Gassner, & Peters, 2012). Reliability of rainfall duration and intensity varies by agro-climatic zone as does soil fertility, market access, temperatures, and suitability of growing conditions for specific crops. Generally, cassava is the primary and staple food crop with beans providing an inexpensive alternative source of protein to meat, dairy, and fish products. Due to these regional variations, this study's agricultural interventions had to be flexible to allow localized tailoring to the diverse set of farmer needs and growing conditions within the sample.¹

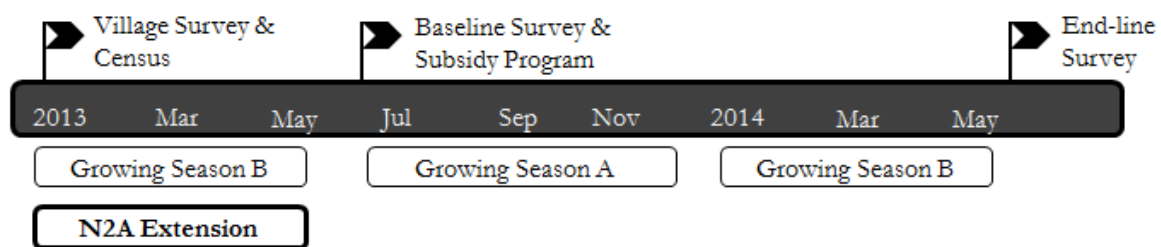
3.2 Intervention

The intervention was run under the umbrella of the N2Africa program. N2Africa program was established in 2009 and operates in 10 countries across sub-Saharan Africa. The over-arching mission of the program is increasing yields, food security, and incomes through agricultural extension work through the production of nitrogen-fixing legumes. The motivation of this mission is that legume production provides a sustainable source of nitrogen to farmer fields, helping combat and even reverse soil fertility degradation, while simultaneously diversifying smallholder farmer production and consumption and offering income diversification through crop sales (Pypers, Sanginga, Kasereka, Walangululu, & Vanlauwe, 2011; Woomer, Huising, Giller, & Et, 2014). This study uses data collected from a community-focused extension and subsidy project undertaken by N2Africa in the eastern DRC arm of the project over the 2013 and 2014 growing seasons.² See **Figure 1** for a timeline of project activities.

¹ See Appendix A1 for details on regional variation within the province.

² The DRC arm of the N2Africa Project consists of a multi-party collaboration between several international and local partners. See appendix A2 for a full list.

Figure 1 Timeline of Research Activities and Regional Agricultural Seasons



Communities were selected to participate in the N2Africa project based on three criteria. First, that the community was located in an area that one of the 6 local implementing partner organizations had previously undertaken projects.³ Second, that each community was accessible by motorized vehicle to ensure successful delivery of all project-distributed resources. Third, a community could not have participated previously in any prior N2Africa projects. Based on funding availability and in cooperation with the local implementing partners, 70 communities were selected to receive the N2Africa agricultural extension program. Of these 70 extension-receiving communities, half were randomly selected to additionally receive an offer of purchasing a package of agricultural inputs at a subsidized price. After this assignment of the initial 70 communities into the two treatment arms, an additional 32 control communities were selected to serve as a control group. Assignment of communities to the control group was therefore not randomized but done by selecting communities comparable to the treated communities. A census in all 102 communities was conducted and the resulting information used to randomly select ten households for surveying within each community.

The 70 treatment communities received the extension training program once all census work had been completed. The extension program began with a period of sensitization on new techniques and inputs with a specific focus on soybean production and its associated inputs. This focus on soybeans is due to the crop's novelty within the region, nutritionally beneficial composition, and its need for inoculation from commercial rhizobia which necessitates information-intensive training for successful production. During this period, local needs and constraints were identified to ensure projects were properly tailored to maximize potential program benefits. Each community elected a 'lead' (or 'master') farmer based on criteria of literacy, property

³ This was to ensure established relationships of trust and cooperation were effectively leveraged.

ownership, extensive farming experience, and access to external sources of advice and supplies of production inputs. Master farmers received training on the application of new techniques and inputs for growing grain legumes, with soybeans receiving special emphasis. In return, contact-farmers were expected to organize regular meetings within their agricultural farmer groups and act as liaisons between the wider community and the implementing partner. All contact-farmers and interested members of their farmer groups were also able to visit research stations at which experimental trials of intercropping soybean with maize or cassava were undertaken.⁴ With the assistance of extension workers, farmer groups would select those options (inputs and practices) most suitable for their respective local conditions. Contact-farmers were then expected to establish their own demonstration plots using their own land. Community farmers interested by the results of demonstration plots could request very small input packages to experiment with privately on their personal fields, however package sizes provided only a very limited level of experimentation.⁵

In the 35 randomly assigned villages to receive an additional subsidy-offer, this subsidized package was made available after all extension service programs concluded in August 2013. Farmers in these randomly selected communities were offered the opportunity to purchase packages of inputs consisting of a combination of improved seeds, in-organic fertilizer, and inoculum at a subsidized rate. Implementing partners custom-designed up to six packages each, where packages contained varying amounts and types of crop seeds, fertilizer, and inoculum of equivalent value in order to best target regional differences and local farmer needs.⁶ Each package was offered to farmers at the equivalent of 26 USD, roughly 75% of package market value. These prices did not take into account transportation costs that were avoided by project-delivery of resources, resulting in additional, but variable, cost-savings for communities. Community-based Local Development Committees (CLDs) were responsible for informing community members of the offer and the associated delayed payback scheme.⁷ CLDs were also tasked with registering applicants, ordering packages, coordinating distribution within the community and collecting final payments.

⁴ These trials were meant to juxtapose improved inputs and management techniques with traditional practices.

⁵ Individuals across both treatment arms had access to these small packages. Throughout the growing season extension workers would visit the communities to provide assistance, advice, and assess on-going results. In some communities field days were organized so that non-participating farmers could also observe project results.

⁶ See appendix A3 for details on package compositions

⁷ Payment was postponed until after harvest was completed. A 500 FC (USD\$0.54) down-payment was required.

As can be seen in Figure 1, eastern DRC is characterized by two growing seasons each year. A long season, A, between July and December and a short season, B, between February and May. Baseline and end-line surveys were conducted immediately following Growing Season B in 2013 and 2014.⁸ Extension services were conducted during Season B of 2013 and subsidized packages were delivered between season B and A of 2013. Every effort was taken to interview the same 10 farmers of each community, although some replacements did occur⁹. The resulting dataset thus has a two-phase panel structure and covers in-depth production indicators as well as socio-economic and demographic details of the household.

4 Empirical Analysis

The analysis phase can be decomposed into three steps. First, relationships between farmer and farm traits with higher farmer output levels are analyzed using an OLS regression of baseline farm-level production and household-specific indicators. Second, project intervention impacts on farmer productivity are estimated using a difference-in-difference model. Fourth, channels through which observed treatment effects work are teased out, focusing primarily on changes in the changes in intercropping patterns, the amount of inputs used in production, value of common bean yields, and use of in-organic fertilizer. I describe each of these empirical analysis models in detail below.

4.1 Farm Characteristics on Farmer Output Empirical Model

Using only baseline data, I regress several farm-level and household covariates on a producer's average plot technical efficiency score using the OLS model

$$Crop\ Value_{ij} = \alpha_0 + \alpha_1 HH\ Head_{ij} + \alpha_2 Farm_{ij} + \alpha_3 IC_{ij} + \varepsilon_i \quad (1)$$

where $Crop\ Value_{ij}$ is the value of a farmer's total output of all produced crops measured first in USD and second in caloric value for farmer i in community j . As the study region is characterized by weak infrastructure and high transaction costs to market, the $HH\ Head_{ij}$ is a vector of household-head characteristics including a social capital score, use of credit in the previous 12 months, age, education level,

⁸ 9 villages participated in the baseline survey and intervention but were not visited in the end-line survey due to logistical constraints and were dropped from the sample.

⁹ No correlation with treatment arms was found. See appendix A5 for attrition and replacement analysis.

gender, migrant status, an indicator for the household having a tin roof, and an indicator for the household head being a member of an agricultural cooperative. $Farm_{ij}$ includes farm-level production decisions including the number of crops produced, number of household members that work on the farm, total farm size, proportion of farm owned by the household, perceived soil quality of the farm, the use of hired labor, and the use of organic or in-organic fertilizer. The term IC_{ij} is a vector of intercropping dummies included to control for the different combinations of plants a farmer decides to produce on each household plot. Standard errors are clustered at the community level.

4.2 Intervention Impact Empirical Model

In total, four outcome variables of interest are used. The first two are both measures of technical efficiency, with one being in terms of USD value of output and the other in terms of caloric content of output. The second two are log transformations of the level values of output, both in terms of USD and caloric content. The first two, technical efficiency, account for inputs used while the second two do not. Using these four outcomes of interest independently, the following difference-in-difference regression analysis is used to determine the impact of each intervention on farm-level productivity.

$$Y_{ijt} = \alpha_0 + \alpha_1 Post + \alpha_1 Post * training_{ij} + \alpha_2 Post * subsidy_{ij} + \alpha_3 X_{ijt} + \alpha_4 IC_{ijt} + \eta_i \quad (2)$$

In which Y_{ijt} is either the USD or calorie valuation of farm output for farmer i from village j in time t , $Post$ is binary indicator for time period with the endline period taking a value of one. Consistent with difference-in-difference approaches, an interaction term for time period ($Post$) with each of the treatment groups is included and the coefficients for these interaction terms is the estimated treatment effect for each of the two interventions respectively. X_{ijt} is a vector of $K \times 1$ farm and producer characteristics for farmer i from village j in time t to control for differences in means that arose between each of the treatment and control groups. This vector includes an indicator on credit use in the last 12 months, indicator of a tin roof, membership of an agricultural cooperative, log transformation of total farm size, percent of farmed land owned, and perceived farm soil quality. Similar to equation (1), IC_{ij} is a vector of intercropping dummies ensure that variations arising due to choice of crop combinations on a given plot are controlled for. Finally, η is the random error term assumed to be identically and independently distributed with distribution $N(0, \sigma_\eta^2)$. Standard errors are clustered at the community level.

4.3 Intervention Mechanism Analysis Empirical Model

To explore channels through which the agricultural interventions had the strongest effects, I isolate treatment impacts on specific changes in production behavior. The four intervention mechanisms focused on are changes in intercropping behavior, input use as measured by an index score, total USD value of common bean yields, and uptake of in-organic fertilizer. These four mechanisms are chosen for their links to content covered in the extension programs. Improved soil fertility management through the production of legumes forms one of the primary pillars of the N2Africa program. Intercropping of legumes with staple crops, improving soil fertility through legume production, as well as benefits and proper use of improved inputs including improved seeds, in-organic fertilizer, and inoculum for soybean were all covered in the extension curriculum. We look at changes in intercropping behavior by looking at the proportion of plots on which a farmer changed the intercropping pattern from the previous year. In collecting data, plots were not assigned unique identifiers, so are not able to be directly linked across data collection waves. As a result we rely on a dependent variable of the proportion of plots that have a different intercropping pattern than observed the previous year. With this construction, models of analysis for panel datasets are inapplicable. Instead, a tobit model is applied bounded between zero using baseline values of covariates.

$$IC\ Proportional\ Change_{ij} = \alpha_0 + \alpha_1 training_{ij} + \alpha_2 subsidy_{ij} + \alpha_3 X_{ij} + \alpha_4 V_i + \eta_i \quad (3)$$

All right hand side variables are equivalent to those described in equation (2) with the addition of V_i , a vector of community fixed effects. Standard errors are also clustered at the community level, similar to equations (2) and (3). For the remaining three mechanisms explored, the difference-in-difference model used in estimating treatment effects on technical efficiency and output value levels is applied.

5 Data

Our study analyses socioeconomic and production data collected from 3,170 plots owned by 953 households within 93 communities throughout the South Kivu province of the Democratic Republic of Congo (DRC).¹⁰ The sample composition for each treatment is described in **Table 1**.

¹⁰10 households were randomly sampled for surveying in each village. Some households participated in the baseline but could not be found for the endline survey collection. Households that were not agricultural producers in the baseline were replaced

Table 1 Sample Size by Treatment

Treatment	Communities	Households	Plots
Control	29	287	997
Extension	33	349	1141
Ext + Subsidy	31	317	1032
Total	93	953	3170

Farmers within the region are primarily subsistence producers but market sales do occur on a regular basis. As the incentives behind production may therefore be a combination of consumption and sale, farmer total crop yields are translated into both US dollar value and caloric value. As crop bundles in the region vary significantly across households, converting all crop yields into single unit valuations allows household total output levels to be compared.¹¹ The reported mean of crop yield values is artificially high due to high noise within the dataset coming from self-reported memory-based crop yields.¹² The noise generated by self-reported crop yields is evident in the high standard deviations associated with both measures of household output. For all analyses, log transformations of yield values are used to minimize the influence of outliers on estimators of treatment effect without subjecting data to subjective cut-off thresholds. For production inputs, farmers produce less than two different crop varieties across their primary four fields. Plots are small, around .38 of a hectare, although this size varies strongly within the sample. Roughly 42% of the sample hired outside labor to work on their farm, and most households have on average of 2 internal members that labor on each plot. On average farmers report their soils as being slightly above moderate fertility levels and while nearly 40% of the sample reports using organic fertilizer on their fields, less than 5% report using inorganic fertilizer to improve soil fertility levels.

with agriculture producing households in the endline. As a result, the total number of households interviewed is greater than 930. See appendix A5 for attrition and replacement correlation with treatment analysis.

¹¹ As households face unique market transaction costs, all valuations were calculated using average reported market prices within the sample for each crop.

¹² Excluding observations above the 95th percentile drops the mean household yields to USD 6,859.53.

Table 2 Summary Statistics of Household Production Outputs and Inputs

	Mean	Std. Dev.
Farm Yield (<i>USD</i>)	29,339.97	37,6511.5
Farm Yield (<i>kJ</i>)	4.23e+07	5.50e+08
Plot Size (<i>Ha</i>)	0.383	1.515
Hired Labor (<i>=1</i>)	0.42	0.494
HH Labor (<i>No. Members</i>)	2.11	1.166
Soil Fertility (<i>1 = Very Infertile; 5 = Very Fertile</i>)	3.355	0.984
Organic fertilizer (<i>=1</i>)	0.398	0.49
In-organic fertilizer (<i>=1</i>)	0.044	0.206
Hired Labor, Organic Fertilizer, and In-organic Fertilizer are measured as indicator variables in which use of the input takes value one. Reported means are the proportion of sample which uses the input.		

Control communities were selected by implementing partners as comparable communities that met the outlined criteria for project participation, but were not chosen for participation. Due to this non-random assignment, I check control group orthogonality against all households in the pooled treatment groups. **Table 3** displays the pre-intervention mean values and standard deviations for the overall sample, the control group, and the pooled treatment group for all indicators used in the technical efficiency analysis. P-values for differences in means between treated communities and control communities are listed in column (5). Between control households and all treated households, no significant differences in means is found for any of the tested covariates.

Participation in the extension-only treatment and extension plus subsidy treatment was assigned through randomization over the originally selected 70 communities. I confirm that random assignment to the two treatment arms successfully achieved orthogonality. **Table 4** reports a difference in means test between the two treatment arms. For all covariates tested, no statistically significant difference in means is detected between the two randomly-assigned treatment groups.

Table 3 Balance Table for Control vs. Treated Households

	(1)	(2)	(3)	(4)	(5)
		Means			
	N	Control	Treated	Overall	(2) vs. (3) p-value
<i>Household Characteristics</i>					
Social Capital Score	748	41.204 (2.085)	41.815 (1.307)	41.627 (1.103)	0.803
Used Credit (past 12 mos.) (=1)	894	0.405 (0.034)	0.394 (0.020)	0.397 (0.017)	0.766
HH Head Age (yrs)	887	47.033 (0.886)	45.571 (0.719)	46.015 (0.572)	0.200
HH Head Edu	892	1.029 (0.058)	0.958 (0.049)	0.980 (0.038)	0.345
HH Head female (=1)	895	0.179 (0.026)	0.145 (0.018)	0.155 (0.015)	0.276
HH Head migrant (=1)	895	0.350 (0.048)	0.386 (0.034)	0.375 (0.028)	0.538
HH Tin Roof (=1)	895	0.526 (0.045)	0.530 (0.032)	0.528 (0.026)	0.938
HH Head Ag. Coop Member (=1)	895	0.113 (0.024)	0.113 (0.014)	0.113 (0.012)	0.988
<i>Production Characteristics</i>					
Total number of crop varieties produced	895	2.026 (0.086)	1.902 (0.047)	1.940 (0.042)	0.206
Total farm size - Log transformed	895	-1.556 (0.126)	-1.377 (0.079)	-1.431 (0.067)	0.229
HH Labor (No. Members)	895	2.569 (0.070)	2.591 (0.073)	2.584 (0.055)	0.830
Hired Labor (=1)	895	0.365 (0.033)	0.380 (0.022)	0.375 (0.018)	0.700
Percentage Farm Owned	888	0.839 (0.029)	0.812 (0.021)	0.820 (0.017)	0.455
Soil Fertility (1 = Very Infertile ; 5 = Very Fertile)	895	3.279 (0.115)	3.216 (0.072)	3.235 (0.061)	0.642
Used Organic Fertilizer (=1)	895	0.449 (0.071)	0.388 (0.049)	0.407 (0.041)	0.482
Used In-organic Fertilizer (=1)	895	0.036 (0.011)	0.034 (0.007)	0.035 (0.006)	0.834

Table 4 Balance Table for Extension vs Extension + Subsidy Households

	(1)	(2)	(3)	(4)
		Means		
	N	Control	Treated	(2) vs. (3) p-value
<i>Household Characteristics</i>				
Social Capital Score	518	41.410 (1.548)	42.226 (2.137)	0.757
Used Credit (past 12 mos.) (=1)	620	0.381 (0.028)	0.407 (0.028)	0.498
HH Head Age (Yrs.)	618	44.802 (1.032)	46.414 (1.000)	0.263
HH Head Edu	618	1.009 (0.061)	0.902 (0.077)	0.276
HH Head female (=1)	621	0.130 (0.021)	0.161 (0.029)	0.386
HH Head migrant (=1)	621	0.356 (0.049)	0.419 (0.046)	0.343
HH Tin Roof (=1)	621	0.523 (0.044)	0.537 (0.047)	0.830
HH Head Ag. Coop Member (=1)	621	0.115 (0.021)	0.111 (0.020)	0.895
<i>Production Characteristics</i>				
Total number of crop varieties produced	621	1.897 (0.068)	1.908 (0.066)	0.908
Total farm size - Log transformed	621	-1.305 (0.110)	-1.454 (0.113)	0.346
HH Labor (No. Members)	621	2.632 (0.113)	2.547 (0.092)	0.560
Hired Labor (=1)	621	0.415 (0.032)	0.342 (0.028)	0.089
Percentage Farm Owned	616	0.825 (0.022)	0.797 (0.036)	0.513
Soil Fertility (1 = Very Infertile; 5 = Very Fertile)	621	3.198 (0.096)	3.236 (0.109)	0.794
Used Organic Fertilizer (=1)	621	0.427 (0.072)	0.346 (0.067)	0.408
Used In-organic Fertilizer (=1)	621	0.028 (0.009)	0.040 (0.011)	0.392

Despite no statistical evidence of systematic differences in households between groups, due to the non-random assignment of households to the control group, I include a selection of relevant dynamic covariates in all regression analyses. Thus for all analyses, time-varying variables that may drive different trends between periods are controlled for.

Table 5 Outcome Variable Means and Standard Deviations by Treatment and for Full Sample

		Means			
	N	Control	Extension	Ext + Subsidy	Overall
<i>Primary Outcomes</i>					
Log of Output Value (USD)	895	8.402 (0.201)	8.412 (0.197)	7.986 (0.255)	8.267 (0.127)
Log of Output value (kJ)	895	14.965 (0.326)	15.130 (0.291)	14.518 (0.401)	14.876 (0.197)
<i>Mechanisms</i>					
Intercropping*	895	0.671 (0.027)	0.625 (0.023)	0.682 (0.021)	0.658 (0.014)
Cultivated Plots	895	1.923 (0.060)	1.839 (0.050)	1.822 (0.060)	1.859 (0.032)
Input Score	895	-0.133 (0.160)	-0.056 (0.153)	-0.001 (0.160)	-0.061 (0.090)
Bean Output (USD)	474	4.050 (0.273)	3.639 (0.334)	3.413 (0.362)	3.683 (0.192)
In-organic Fertilizer	895	0.036 (0.011)	0.028 (0.009)	0.040 (0.011)	0.035 (0.006)

* Measure of change in intercropping behaviour between time periods not a baseline value

Bold results indicates a significant difference in means between Extension-only and Extension

+ Subsidy treatment groups at the 10% significance level

Summary statistics and balance between treatments of baseline values for outcome variables are displayed in **Table 5**. The exception of baseline values being used is made with the outcome variable “Intercropping”, which is itself a variable measuring change between time periods. This is the only variable with significant difference in means between the two treated groups but does not suggest a systemic difference in treatment groups pre-intervention. All other variables are balanced across treatments in the baseline period suggesting that households across groups were not systemically different prior to receiving the study interventions for all outcomes measured.

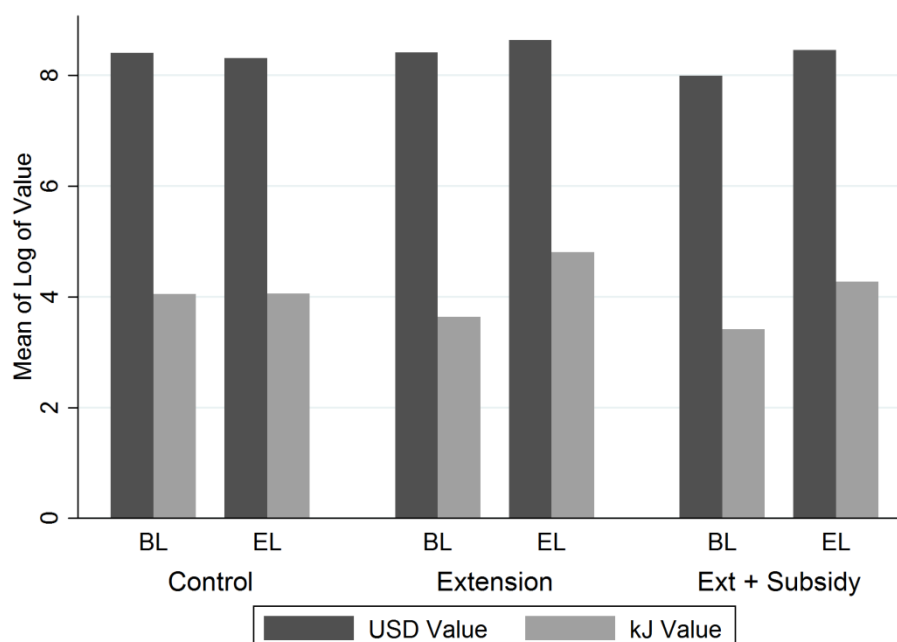
6 Results

6.1 Farmers and Output Levels

I analyze the relationships between household head socio-economic indicators and input resources with observed total output values. While the control group experiences a marginal decrease in output value, for both USD and caloric content measures, both intervention groups experience slight growth in output between pre- and post-intervention time periods (**Figure 2**). Since prices used in calculating USD value

were constant between time periods, this growth is attributed entirely to higher output levels rather than price effects.

Figure 2 Farmer Production in Dollar and Caloric Value by Treatment and Time Period



Additionally, I evaluate potential change in input use due to the complementary nature of production factors. Using principal component analysis on the major production function inputs of land, labor, soil fertility, and fertilizer in the form of either organic or in-organic, a single input score is calculated for each farmer. **Figure 3** plots this score by treatment and time period to evaluate trends in input use. Input use increases dramatically in the endline time period for all treatments, suggesting that increased input intensity is driven by changes in factor market prices rather than from extension or subsidy interventions.

Next I explore what observed farm and farmer characteristics correlate with higher farm output levels. I regress several baseline characteristics on both the dollar and caloric value of farmer crop output. Results of these OLS regressions are presented in **Table 5**. Household characteristics appear to have low correlation with overall farm yields. Agricultural cooperative membership is positively and weakly significant in its relationship with the USD value of farm output but this significance is not evident with caloric values of yield.

This result may be explained by the role of cooperatives in marketing products, resulting in farmers that place higher emphasis on profit-maximization over subsistence production being more likely to both participate and benefit from cooperative membership.

Figure 3 Farm-level Input Use by Time Period and Treatment

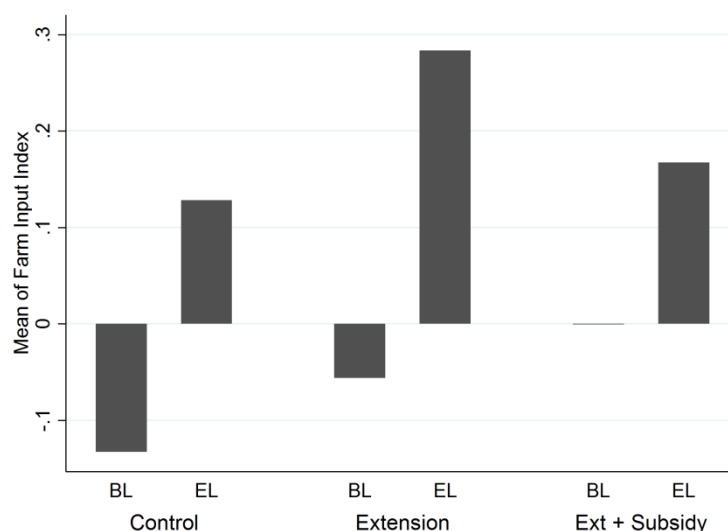


Table 5 Outcome Variable Means and Standard Deviations by Treatment and for Full Sample

		Means			
	N	Control	Extension	Ext + Subsidy	Overall
<i>Primary Outcomes</i>					
Log of Output Value (USD)	895	8.402 (0.201)	8.412 (0.197)	7.986 (0.255)	8.267 (0.127)
Log of Output value (kJ)	895	14.965 (0.326)	15.130 (0.291)	14.518 (0.401)	14.876 (0.197)
<i>Mechanisms</i>					
Intercropping*	895	0.671 (0.027)	0.625 (0.023)	0.682 (0.021)	0.658 (0.014)
Cultivated Plots	895	1.923 (0.060)	1.839 (0.050)	1.822 (0.060)	1.859 (0.032)
Input Score	895	-0.133 (0.160)	-0.056 (0.153)	-0.001 (0.160)	-0.061 (0.090)
Bean Output (USD)	474	4.050 (0.273)	3.639 (0.334)	3.413 (0.362)	3.683 (0.192)
In-organic Fertilizer	895	0.036 (0.011)	0.028 (0.009)	0.040 (0.011)	0.035 (0.006)

* Measure of change in intercropping behaviour between time periods not a baseline value

Bold results indicates a significant difference in means between Extension-only and Extension

+ Subsidy treatment groups at the 10% significance level

Far more relationships between production characteristics with productivity are found. More diverse crop bundles are estimated to be more likely associated with higher economic returns and while a positive relationship is estimated with consumption measures of output, no statistical significance is found. Given that household members provide the primary source of farm labor within the region, it is unsurprising that larger households are correlated with higher output levels. Finally we observe similar phenomenon observed in many developing countries of larger farms being associated with lower output levels, *ceteris paribus*. No relationship is found between inputs of stronger farm property rights, perceived soil fertility, use of hired labor, or use of any fertilizer with household production levels. As in-organic fertilizer use within the sample is extremely low (less than 5% of the sample) the sample may suffer from insufficient power for accurately reflecting relationships between in-organic fertilizer and crop yields.

Overall, crop diversification, more intensive farming, and household labor are all positively associated with higher crop output valuation. While crop diversification and intensive farming are more manipulable short term variables, adjustments in labor inputs is reliant on the presence and strength of local labor markets. Neither crop diversification nor intensive farming are primary foci of the extension services studied here, however both are touched on indirectly through the introduction of soybean production and improved farm management strategies (including planting in rows and soil fertility management) discussed in the agricultural trainings.

Table 6 Farmer and Farm Characteristics with Household Production Levels

VARIABLES	(1) Log Yield (USD)	(2) Log Yield (kJ)
<i>Household Characteristics</i>		
Social Capital Score	0.007 (0.005)	0.012 (0.009)
Used Credit (past 12 mos.)	0.152 (0.267)	0.426 (0.443)
HH Head Age (Yrs.)	-0.007 (0.009)	-0.017 (0.015)
<i>Highest Education Attained</i>		
Primary	0.181 (0.388)	0.376 (0.646)
Secondary	0.128 (0.358)	0.213 (0.569)
Tertiary	0.640 (0.478)	0.898 (0.866)
Female (=1)	0.315 (0.368)	0.770 (0.588)
Migrant (=1)	0.268 (0.319)	0.329 (0.545)
Tin Roof (=1)	0.330 (0.309)	0.460 (0.522)
Ag. Coop Member (=1)	0.571* (0.304)	0.739 (0.501)
<i>Production Characteristics</i>		
Crop diversity (No. Crops)	0.582* (0.302)	0.779 (0.497)
HH Labour (No. Members)	0.217** (0.104)	0.347** (0.173)
Farm Size (Log of Ha)	-0.773*** (0.087)	-0.810*** (0.136)
Percentage Farm Owned	0.400 (0.410)	0.642 (0.657)
Soil Fertility (1 = Very Infertile; 5 = Very Fertile)	0.168 (0.143)	0.224 (0.238)
Hired Labor (=1)	0.440 (0.272)	0.511 (0.446)
Organic Fertilizer (=1)	0.077 (0.453)	0.053 (0.746)
In-organic fertilizer (=1)	-0.125 (0.827)	-0.466 (1.347)
Constant	2.993** (1.249)	8.520*** (2.129)
Observations	582	582
R-squared	0.401	0.335

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.2 Intervention Impacts

Measuring the treatment effect of extension only and extension combined with a subsidy input-package offer a difference-in-difference fixed effects model is run on the two measures of farmer output. Results for both regressions are presented in **Table 7**. Each regression includes time-varying control variables but only

coefficients of estimated treatment effects, time trends, and means (constant) are displayed. Additionally, p-values for F-tests testing significant differences between the treatment effects of the two interventions are included in rows at the bottom of **Table 7**.

Table 7 Treatment Effect Estimates on Technical Efficiency and Total Farm Output Value

VARIABLES	(3) Log USD	(4) Log kJ
Extension * Post	0.245 (0.377)	0.375 (0.604)
Extension+ Subsidy * Post	0.956** (0.376)	1.450** (0.578)
Post-Intervention	-0.233 (0.255)	-0.369 (0.385)
Constant	5.073*** (0.610)	10.823*** (1.003)
Observations	1,650	1,650
R-squared	0.165	0.099
Number of households	951	951
<i>P-Value of F-Test</i>		
Extension * Post v. Ext + Sub * Post	0.0681	0.0868

All models include control variables of household credit use in the last 12 months, roof material of house, membership in an agricultural cooperative, log transformation of total farm size (Ha), proportion of farm owned by household, and perceived farm soil quality (scale of 1 = very infertile to 5 = very fertile).

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

A positive but insignificant impact of the extension only intervention is observed, ranging between 24% and 37.5% increase in USD and caloric measures of output, respectively. For the combined treatment of extension offered with input subsidization, large and significant gains in farm output are estimated for both economic and consumption measures of total farm output value. Additionally, for both regressions the difference in estimated intention to treat effects between treated groups is statistically distinguishable at the 10% significance level. These results would indicate that targeting both farmer knowledge and associated transaction costs simultaneously generates larger returns compared to targeting knowledge independently.

6.3 Mechanisms Analysis

To better explore potential changes in farmer decision-making I explore four mechanisms through which observed increases in output value in the subsidy treatment group could be operating. Results for each mechanism is displayed in **Table 8**. No significant results are observed for the extension only treatment

arm on any of the outcomes analyzed and no effects from either treatment arm are found on common bean yields.

Table 8 Results for Impacts on Secondary Outcomes: Intercropping, Input Use Score, Bean Yields (Kg/Ha), and Use of Fertilizer

VARIABLES	(1) Intercrop	(2) Bean Yield	(3) Inputs Score	(4) Fertilizer
Extension Only * Post		0.054 (0.044)	-0.014 (0.097)	0.003 (0.022)
Ext + Subsidy * Post		0.023 (0.047)	-0.156** (0.071)	0.058** (0.029)
t		-0.026 (0.035)	0.136*** (0.051)	0.003 (0.016)
Extension Only	-0.085 (0.069)			
Ext + Subsidy	0.031 (0.066)			
Constant	0.694*** (0.125)	0.273*** (0.053)	-1.747*** (0.142)	0.031 (0.038)
Observations	887	843	1,650	1,650
R-squared		0.078	0.505	0.023
Number of Households		613	951	951
Sigma	0.623*** (0.031)			
<i>P-value of F-test:</i>				
Ext * Post vs. Ext + Sub * Post	0.0519	0.450	0.136	0.0458
All models include control variables of household credit use in the last 12 months, roof material of house, membership in an agricultural cooperative, log transformation of total farm size (Ha), proportion of farm owned by household, and perceived farm soil quality (Likert scale of 1 to 5).				

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The first two columns of **Table 8** focus on production relationships of output-output, exploring the crop activity choices of households within the sample. The change in the proportion of plots on which a farmer changed their intercropping behavior is analyzed using a tobit model (Column 1). Neither treatment arm is statistically distinguishable from the control group, however they are statistically difference from one another. Farmers in the extension only group appear less willing to adjust intercropping behavior while farmers in the combined extension and subsidy group demonstrate a higher likelihood of changing their intercropping behavior between time periods. This result may indicate that farmer experimentation with improved seeds from the subsidized package in the prior season were more likely to adjust their intercropping behavior in reallocating resources to new outputs. However absent more detailed information on adoption of improved seeds at the individual level, this outcome cannot be confirmed concretely.

Column (2) reports estimates of the impact of the treatment arms on bean yields, for which observable relationship is found. This result indicates that increases in total farm output observed for the combined extension with subsidized inputs treatment arm were not driven by beans but potentially across crop portfolios. Common beans are an existing crop within the region and are unlikely to qualify as a new innovation, making additional experimentation unlikely to influence farmer adoption decisions.

In terms of changes in input use, two specific mechanisms are analyzed in Column (3) and (4) of **Table 8**. Column (3) reports results of changes in overall input use as measured by an index of all production function variables. Column (4) analyzes change in in-organic fertilizer use. Both indicators are found to have a significant relationship with the combined extension and subsidy treatment. While an overall increase in input use is estimated to have occurred in the post-intervention time period, households in the combined extension with subsidy treatment group had lower average input bundle scores compared to control and extension-only households. However producers in the extension + subsidy treatment arm are nearly 6% more likely to adopt in-organic fertilizer on their fields. This rise in in-organic fertilizer suggests that reductions in overall input are sufficiently large as to offset this increased use of in-organic fertilizer. Furthermore, the decrease in inputs is contributing to the overall increase in farmer output levels observed for the subsidy treatment arm in Table 7. These two results combined would indicate that households are benefiting from synergistic relationships between inputs that may not be being realized by households in the control and extension-only groups.

7 Conclusion

Incomplete information, limited financial endowments of farmers, and highly variable market transaction costs constrains proper use of and access to productivity enhancing inputs, contributing to the persistence of low yields for African agriculture household producers. In particular, information gaps on the returns and use of improved inputs has been found to be a major obstacle to widespread adoption within sub-Saharan Africa. Programs providing agricultural extension have been proposed as tools to target knowledge gaps. These programs rely heavily on social learning and diffusion of information through community networks. However information biases that result from observed rather than personally experienced engagement with new innovations pose a threat to the effectiveness of social learning-based extension

designs. Providing reduced-cost opportunities to experiment directly with improved inputs offers a second policy tool that targets information gaps with a learning-by-doing approach. I test the effectiveness of both an independent agricultural extension program and a combined agricultural extension and input subsidy program on improving farmer production levels within eastern Democratic Republic of Congo (DRC). Due to the non-separability of production and consumption decisions made by agricultural households, I estimated total yields both in economic (USD) and consumption (caloric content) values.

I address three important areas within existing literature. First, I analyze what farmer traits are correlated with higher farmer crop output within the region using an OLS regression of pre-intervention farm and household characteristics. Second, I estimate intervention impacts using a difference-in-difference approach. Third, I test channels through which the interventions may be generating their estimated treatment effects on household crop yields. These channels include changes in intercropping patterns on farm plots, changes in input use between time periods, common bean yields (Kg/Ha) and farmer use of in-organic fertilizer.

I find that farmer output levels are more closely associated with production decisions rather than household characteristics. The positive relationship between crop diversification and farmer output suggests that farmers face high uncertainty stemming from fluctuating market prices making diverse crop portfolios an economic risk-mitigation strategy. Decisions on the intensity with which land is farmed and considerations of land tenure are additional areas agricultural development programs can explore in efforts to increase farmer production levels.

Difference-in-difference analysis on the effect of each treatment arm on total household output estimated positive effects for both treatment groups, although only the extension-plus-subsidy intervention had statistical significance. These results indicate that knowledge gaps do present a constraint to agricultural development within the region, and targeting these information gaps with social learning and reduced-cost experimentation yields larger effects on household output than social learning programs alone. As no subsidy-only treatment arm was undertaken, it cannot be determined whether the observed impact is the result purely of the greater access to experimentation, whether synergistic benefits arose from reinforcement between information gleaned via social learning and independent experimentation.

The subsidized input packages were a one-off shift in transaction costs so the only long-run change generated by the intervention is in addressing uncertainty around output variability within the technology adoption decision. This result therefore indicates that experimentation offers a stronger mechanism for adjusting household production functions, and perceived returns to technology adoption, compared to solely information provision. As our empirical model estimates the intention to treat effect, it is not possible to decompose experimentation into direct learning-by-doing (i.e. treatment uptake) compared to observation (i.e. observers of adopters). It is likely that our observed intention to treat effect is a combination of both the direct and indirect benefits of farmer experimentation.

The coefficient estimate on the extension-only impact on productivity is large and positive but was not found to be statistically significant. The coefficient on the combined extension and subsidy group is only weakly significant, and it cannot be ruled out that the two projects generate relatively comparable results. However the cost of implementation for the provision of subsidized inputs is substantial and very well may eclipse the relative gains in production outputs that farmers are able to realize given the existing market institutions and infrastructure within the region. A full cost-effectiveness analysis is beyond the scope of this study, but warrants future research before committing to the conclusion that combined extension and subsidy programs is a stronger policy tool for agricultural development within the region.

Our study was conducted in an impoverished region of one of the least developed countries in Africa. The area is subject to persistent instability and characterized by high variation in topography and climate. Farmer output levels within the region are low, even by comparison to neighboring countries, suggesting large gains are available for even limited agricultural interventions. A similar project model may not be successful in countries with more developed agricultural sectors and higher market integration. Further research is required on how information constraints to agriculture evolve with better road and ICT infrastructure, higher market integration, and more complete markets.

In addition, our intervention was implemented by six independent partners. While standardized protocols for project implementation were established, inevitable variation in partner structures, beneficiary relationships, and processes existed. While this multi-partner framework more accurately reflects real world development practices, these variations in regional traits and partner implementation variations make

generalizability of results limited to comparable project structures implemented under similar local conditions.

References

- Agyei-Holmes, A., Ayerakwa, H. M., Osei, R. D., & Osei-Akoto, I. (2011). *Training and Farmer Productivity : An Evaluation using RCT for the MCA-Ghana Programme*.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2010). Preference heterogeneity in experiments: Comparing the field and laboratory. *Journal of Economic Behavior and Organization*, 73(2), 209–224. <https://doi.org/10.1016/j.jebo.2009.09.006>
- Anderson, J. R., & Feder, G. (2007). Chapter 44 Agricultural Extension. *Handbook of Agricultural Economics*, 3(6), 2343–2378. [https://doi.org/10.1016/S1574-0072\(06\)03044-1](https://doi.org/10.1016/S1574-0072(06)03044-1)
- Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America*, 106(51), 21544–21549. <https://doi.org/10.1073/pnas.0908800106>
- Aral, S., Muchnik, L. E. V., & Sundararajan, A. (2013). Engineering social contagions: Optimal network seeding in the presence of homophily. *Network Science*, 1(2), 125–153. <https://doi.org/10.1017/nws.2013.6>
- Aral, S., & Walker, D. (2012). Identifying Influential and Susceptible Members of Social Networks. *Science*, 337, 337–341. <https://doi.org/10.1126/science.1215842>
- Bailard, C. S. (2010). Mobile Phone Diffusion and Corruption in Africa. *Liberation Technologies Seminar Series*. <https://doi.org/10.1080/10584600903053684>
- Bandiera, O., & Rasul, I. (2006). SOCIAL NETWORKS AND TECHNOLOGY ADOPTION. *The Economic*, 116(514), 869–902.
- Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013a). The diffusion of microfinance. *Science*, 341, 363–370. <https://doi.org/10.1126/science.1236498>
- Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013b). The diffusion of microfinance. *Science*, 341(6144), 1236498. <https://doi.org/10.1126/science.1236498>
- Banerjee, A. V., & Duflo, E. (2009). The Experimental Approach to Development Economics. *Annual Reviews of Economics*, 1, 151–178. <https://doi.org/10.1146/annurev.economics.050708.143235>
- Barrett, C. B., & Carter, M. R. (2010). The power and pitfalls of experiments in development economics: Some non-random reflections. *Applied Economic Perspectives and Policy*, 32(4), 515–548. <https://doi.org/10.1093/aep/pqp023>
- Beaman, L., Benyishay, A., William, C., & Mobarak, A. M. (2015). *Can Network Theory based Targeting Increase Technology Adoption ?*
- Benjamin, D. (1992). Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models. *Econometrica*, 60(2), 287–322.
- BenYishay, A., & Mobarak, A. M. (2014). Social Learning and Communication. *NBER Working Paper*, (20139). <https://doi.org/10.3386/w20139>
- Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist*, 37(3).
- Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, 2(1), 113–120.
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29, 555–564. <https://doi.org/10.1016/j.socnet.2007.04.002>
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27, 55–71. <https://doi.org/10.1016/j.socnet.2004.11.008>
- Borgatti, S. P., Carley, K. M., & Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, 28, 124–136. <https://doi.org/10.1016/j.socnet.2005.05.001>
- Bulte, E., Leuvel, K., Nillesen, E. E. M., Pieters, J., & Voors, M. (2015). *Farm Households in Eastern Congo, Baseline Survey Report*.
- Bursztyn, L., Ederer, F., Ferman, B., & Yuchtman, N. (2014). Understanding Mechanisms Underlying Peer Effects: Evidence from a Field Experiment on Financial Decisions. *Econometrica*, 82(4), 1273–1301.
- Cai, J. (2012). Social Networks and the Decision to Insure : Evidence from Randomized Experiments in China *.
- Cai, J., Janvry, A. De, & Sadoulet, E. (2015). Social Networks and the Decision to Insure (Supplementary Figures and Tables). *American Economic Journal: Applied Economics*, 7(2), 81–108. <https://doi.org/10.2139/ssrn.2161686>
- Carter, M. R., Laajaj, R., & Yang, D. (2014). Subsidies and the persistence of technology adoption: Field experimental evidence from Mozambique. *Nber, w20465*. <https://doi.org/10.3386/w20465>
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329, 1194–1197. <https://doi.org/10.1126/science.1185231>
- Centola, D. M., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3), 702–734. <https://doi.org/10.1086/521848>
- Cerdán-Infantes, P. (2008). The impact of agricultural extension services: The case of grape production in Argentina. *OVE Working Papers*. Retrieved from <http://www.iadb.org/intal/intalcdi/PE/2009/03159.pdf>
- CGIAR. (2014). *Project Cialca DR Congo, Rwanda, Burundi*.

- Chami, G. F., Ahnert, S. E., Voors, M. J., & Kontoleon, A. A. (2014). Social network analysis predicts health behaviours and self-reported health in African villages. *PLoS ONE*, 9(7). <https://doi.org/10.1371/journal.pone.0103500>
- Chandrasekhar, A. G., & Lewis, R. (2010). *Econometrics of sampled networks*. *Networks*. <https://doi.org/10.1017/CBO9781107415324.004>
- Choi, H., Kim, S.-H., & Lee, J. (2010). Role of network structure and network effects in diffusion of innovations. *Industrial Marketing Management*, 39(1), 170–177. <https://doi.org/10.1016/j.indmarman.2008.08.006>
- Christiaensen, L., Pan, L., & Wang, S. (2010). *Drivers of Poverty Reduction in Lagging Regions: Evidence from rural Western China* (No. 2010,35). <https://doi.org/10.1007/s10273-011-1262-2>
- Chuang, Y., & Schechter, L. (2000). Social Networks in Developing Countries. *Annual Review of Resource Economics*, (7). <https://doi.org/10.1146/annurev-resource-100814-125123>
- CIALCA. (2007). *Progress Report December 2007. Presentations*.
- Cilliers, J., Kasirye, I., Leaver, C., Serneels, P., & Zeitlin, A. (2014). *Improving Teacher Attendance Using a Locally Managed Monitoring Scheme : Evidence From Ugandan Primary Schools* (Policy Note No. 14/0189).
- Coghlan, B., Ngoy, P., Mulumba, F., Hardy, C., Bemo, V. N., Stewart, T., ... Brennan, R. (2007). *Mortality in the Democratic Republic of Congo: An Ongoing Crisis*. *International Rescue Committee*. Retrieved from http://www.rescue.org/sites/default/files/migrated/resources/2007/2006-7_congomortalitysurvey.pdf
- Conley, B. T. G., & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *The American Economic Review*, 100(1), 35–69.
- Crawford, E. W., & Kelly, V. a. (2006). *Alternative Approaches for Promoting Fertilizer Use in Africa* (Agriculture and Rural Development Discussion Paper No. 22).
- Cunguara, B., & Moder, K. (2011). Is Agricultural Extension Helping the Poor? Evidence from Rural Mozambique. *Journal of African Economics*, 20(4), 562–595.
- Dercon, S., & Christiaensen, L. (2008). *Consumption Risk, Technology Adoption, and Poverty Traps: Evidence from Ethiopia*. (World Economy & Finance Working Paper Series No. WEF 0035).
- Dixit, A., & Pindyck, R. (1994). *Investment Under Uncertainty*. *Investment Under Uncertainty*. Princeton University Press. <https://doi.org/10.2307/1909571>
- Druilhe, Z., & Barreiro-hurlé, J. (2012). *Fertilizer subsidies in sub-Saharan Africa* (ESA Working Paper No. 12–4). Retrieved from www.fao.org/economic/esa
- Duflo, E. (2006). Field Experiments in Development Economics 1. *Econometric Society Monographs*, 42(12), 1–30. Retrieved from <http://qed.econ.queensu.ca/pub/students/phds/khana/econ239/readings/duflo.pdf>
- Duflo, E., Kremer, M., & Robinson, J. (2011). How high are rates of return to fertilizer? Evidence from field experiments in Kenya. *The American Economic Review*, 101(6), 2350–2390. <https://doi.org/10.1257/aer.98.2.482>
- Dutta, S., Lanvin, B., & Paua, F. (2004). *The Global Information Technology Report 2003-2004 Towards an Equitable Information Society*.
- Ellis, F. (1993). *Peasant Economics: farm household in agrarian development* (2nd ed.). Cambridge University Press.
- Emerick, K., Janvry, A. De, Sadoulet, E., & Dar, M. H. (2016). *Identifying early adopters , enhancing learning , and the diffusion of agricultural technology*.
- FAO. (2000). *Fertilizers and Their Use*. *Tennessee University Extension*. FAO Rome. [https://doi.org/10.1016/S0022-3913\(12\)00047-9](https://doi.org/10.1016/S0022-3913(12)00047-9)
- FAO. (2009). *The state of food and agriculture: Livestock in balance*. *Food and Agriculture Organisation of the United Nations*. [https://doi.org/10.1016/S0140-6736\(75\)92740-3](https://doi.org/10.1016/S0140-6736(75)92740-3)
- FAO. (2015). *FAO Statistical Pocketbook 2015*. *Food and Agriculture Organization of the United Nations*. <https://doi.org/978-92-5-108802-9>
- Feder, G., Anderson, J. R., Birner, R., & Deininger, K. (2010). Promises and Realities of Community-Based Agricultural Extension. *Community, Market and State in Development*, (March), 187–210. https://doi.org/10.1057/9780230295018_12
- Feder, G., Murgai, R., & Quizon, J. B. (2003). *Sending Farmers Back to School: The Impact of Farmer Field Schools in Indonesia* (No. 3022). *Policy Research Working Paper*. Retrieved from <http://econ.worldbank.org>
- Feder, G., & Slade, R. (1984). Information Acquisition and the Adoption of New Technology. *American Journal of Agricultural Economics*, 66(3), 312–320. <https://doi.org/10.1287/mnsc.31.11.1372>
- Feder, G., & Umali, D. L. (1993). The Adoption of Agricultural Innovations. *Technological Forecasting and Social Change*, 43, 215–239.
- Ferment, a. M., van Asten, P. J. a, Tittone, P., van Wijk, M. T., & Giller, K. E. (2009). Closing the cassava yield gap: An analysis from smallholder farms in East Africa. *Field Crops Research*, 112(1), 24–36. <https://doi.org/10.1016/j.fcr.2009.01.009>
- Foltz, J. D., Aldana, U. T., & Laris, P. (2012). *The Sabel's Silent Maize Revolution: Analyzing Maize Productivity in Mali at the Farm Level*. *National Bureau of Economic Research Working Paper Series* (Vol. No. 17801). <https://doi.org/10.3386/w17801>
- Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of Technology Adoption. *Annual Review of Economics*, 2. <https://doi.org/10.1146/annurev.economics.102308.124433>
- Gneezy, U., & Imas, A. (2017). Lab in the Field: Measuring Preferences in the Wild. *Handbook of Field Experiments*, *Forthcoming*, 1–40. <https://doi.org/10.1017/CBO9781107415324.004>

- Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., ... Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science*, 327, 812–818. <https://doi.org/10.1126/science.1185383>
- Grossman, G., Humphreys, M., & Sacramone-Lutz, G. (2014). “I wld like u WMP to extend electricity 2 our village”: On Information Technology and Interest Articulation. *American Political Science Review*, 108(3), 688–705. <https://doi.org/10.1017/S0003055414000331>
- Haider, R., Ashworth, A., Kabir, I., & Huttly, S. R. (2000). Effect of Community-Based Peer Counsellors on Exclusive Breastfeeding Practices in Dhaka, Bangladesh: A Randomised Controlled Trial. *Lancet*, 356(9242), 1643–1647.
- Hanna, R., Schwartzstein, J., Benjamin, D., Gentzkow, M., Halac, M., Newman, A., & Ogden, T. (2014). Learning Through Noticing: Theory and Experimental Evidence in Farming.
- Hardy, M., & Mccasland, J. (2015). *It Takes Two : Experimental Evidence on the Determinants of Technology Diffusion*.
- Harrison, G. W., & List, J. a. (2004). Field Experiments. *Journal of Economic Literature*, 42(4), 1009–1055.
- Hiebert, D. L. (1974). Risk, Learning, and the Adoption of Fertilizer Responsive Seed Varieties. *American Journal of Agricultural Economics*, 56(4), 764–768.
- Irz, X., Lin, L., Thirtle, C., & Wiggins, S. (2001). Agricultural Productivity Growth and Poverty Alleviation. *Development Policy Review*, 19(4), 449–466. <https://doi.org/10.1111/1467-7679.00144>
- Janvry, A. De, Fafchamps, M., & Sadoulet, E. (1991). Peasant Household Behaviour with Missing Markets : Some Paradoxes Explained. *The Economic Journal*, 101(409), 1400–1417.
- Just, R. E., & Zilberman, D. (1983). Stochastic Structure , Farm Size and Technology Adoption in Developing. *Oxford Economic Papers*, 35(2), 307–328.
- Kenzior, J., Zibika, J. P., & Voors, M. (2015). *Social relationships, local institutions, and the diffusion of improved variety seed and field management techniques in rural communities: six case studies in South Kivu, DRC*.
- Key, N., Sadoulet, E., & Janvry, A. De. (2000). Transactions Costs and Agricultural Household Supply Response. *American Journal of Agricultural Economics*, 82(2), 245–259. <https://doi.org/10.1111/0002-9092.00022>
- Kim, D. A., Hwong, A. R., Stafford, D., Hughes, D. A., O'Malley, A. J., Fowler, J. H., & Christakis, N. A. (2015). Social network targeting to maximise population behaviour change: A cluster randomised controlled trial. *The Lancet*, 386, 145–153. [https://doi.org/10.1016/S0140-6736\(15\)60095-2](https://doi.org/10.1016/S0140-6736(15)60095-2)
- Kimbrough, E., Smith, V. L., & Wilson, B. J. (2008). Historical Property Rights, Sociality, and the Emergence of Impersonal Exchange in Long-distance Trade. *American Economic Review*, 98(3).
- Kincaid, D. L. (2000). Social Networks, ideation, and contraceptive behavior in Bangladesh: A longitudinal analysis. *Social Science & Medicine*, 50(2).
- Kondylis, F., Mueller, V., & Zhu, S. J. (2014). *Seeing is believing? Evidence from an extension network experiment* (Policy Research Working Paper No. 7000).
- Lambrecht, I., Vanlauwe, B., & Maertens, M. (2016). Agricultural extension in Eastern Democratic Republic of Congo: does gender matter? *European Review of Agricultural Economics*, jbv039. <https://doi.org/10.1093/erae/jbv039>
- Levitt, S. D., & List, J. A. (2007). About the Real World ? *Journal of Economic Perspectives*, 21(2), 153–174.
- Maass, B. L., Musale, D. K., Chiuri, W. L., Gassner, A., & Peters, M. (2012). Challenges and opportunities for smallholder livestock production in post-conflict South Kivu, eastern DR Congo. *Tropical Animal Health and Production*, 44(6), 1221–1232. <https://doi.org/10.1007/s11250-011-0061-5>
- Maertens, A., & Barrett, C. B. (2012). Measuring social networks' effects on agricultural technology adoption. *American Journal of Agricultural Economics*, (649330).
- Marenja, P. P., & Barrett, C. B. (2009). State-conditional fertilizer yield response on Western Kenyan Farms. *American Journal of Agricultural Economics*, 91(4), 991–1006. <https://doi.org/10.1111/j.1467-8276.2009.01313.x>
- Morris, M., Kelly, V. a, Kopicki, R. J., & Byerlee, D. (2007). *Fertilizer uses in African Agriculture: Lessons Learned and Good Practice Guidelines*. Washington D.C.: The World Bank.
- Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *The Quarterly Journal of Economics*, (May), 549–599. <https://doi.org/10.1162/003355303321675455>
- Nisrane, F., Berhane, G., Asrat, S., Getachew, G., Taffesse, A. S., & Hoddinott, J. (2011). *Sources of Inefficiency and Growth in Agricultural Output in Subsistence Agriculture : A Stochastic Frontier Analysis* (No. 19). *Ethiopia Strategy Support Program (ESSP II)*. Retrieved from WWW.ifpri.org
- Nourani, V. (2016). *Social Network Effects of Technology Adoption : Investing with Family , Learning from Friends & Reacting to Acquaintances*.
- Olken, B. A. (2006). Corruption and the costs of redistribution: Micro evidence from Indonesia. *Journal of Public Economics*, 90(4–5), 853–870. <https://doi.org/10.1016/j.jpubeco.2005.05.004>
- Oster, E., & Thornton, R. (2012). Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-Up. *Journal of the European Economic Association*, 10(6), 1263–1293. <https://doi.org/10.1111/j.1542-4774.2012.01090.x>
- Pentland, A. (2014). *Social Physics: How good ideas spread-the lessons from a new science*. Penguin.
- Pypers, P., Sanginga, J. M., Kasereka, B., Walangululu, M., & Vanlauwe, B. (2011). Increased productivity through integrated soil fertility management in cassava-legume intercropping systems in the highlands of Sud-Kivu, DR Congo. *Field Crops Research*, 120(1), 76–85. <https://doi.org/10.1016/j.fcr.2010.09.004>
- Reinikka, R., & Svensson, J. (2005). FIGHTING CORRUPTION TO IMPROVE SCHOOLING : EVIDENCE

- FROM A NEWSPAPER CAMPAIGN IN UGANDA. *Journal of the European Economic Association*, 3(2), 259–267.
- Roe, B., & Just, D. (2009). Internal and External Validity in Economics Research: Tradeoffs between Experiments, Field Experiments, Natural Experiments, and Field Data. *American Journal of Agricultural Economics*, 91(5), 1266–1271.
- Rogers, E. M. (1995). *Diffusion of Innovations*.
- Rosegrant, M. W., & Cline, S. a. (2003). Global Food Security: Challenges and Policies. *Science*, 302, 1917–1919. <https://doi.org/10.1126/science.1092958>
- Simpson, B. M., Franzel, S., Degrande, A., Kundhlande, G., & Tsafack, S. (2015). Farmer-to-Farmer Extension: Issues in Planning and Implementation. *MEAS Technical Note*, (May).
- Smith, A. (1759). *The Theory of Moral Sentiments*.
- Smith, R. D. (2006). It's not just what you do, it's the way that you do it: The effect of different payment card formats and survey administration on willingness to pay for health gain. *Health Economics*, 15(3), 281–293. <https://doi.org/10.1002/hec.1055>
- Taye, H. (2013). Evaluating the impact of agricultural extension programmes in sub-Saharan Africa: Challenges and prospects. *African Evaluation Journal*, 1(1), 1–9. <https://doi.org/10.4102/aej.v1i1.19>
- Taylor, J. E., & Adelman, I. (2003). Agricultural Household Models : Genesis , Evolution , and Extensions. *Review of Economics of the Household*, 1(1).
- Tittonell, P., & Giller, K. E. (2013). When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research*, 143, 76–90. <https://doi.org/10.1016/j.fcr.2012.10.007>
- UNDP. (2013). *2013 Human Development Report: The Rise of the South: Human Progress in a Diverse World*.
- Valente, T. W. (1996). Network models of the diffusion of innovations. *Computational & Mathematical Organization Theory*, 2(2), 163–164.
- Valente, T. W. (2012). Network Interventions. *Science*, 337(6090), 49–53. <https://doi.org/10.1126/science.1217330>
- Valente, T. W., & Davis, R. L. (1999). Accelerating the Diffusion of Innovations using Opinion Leaders. *The ANNALS of the American Academy of Political and Social Science*. <https://doi.org/10.1177/0002716299566001005>
- Valente, T. W., & Pumpuang, P. (2007). Identifying Opinion Leaders to Promote Behavior Change. *Health Education & Behavior*, 34(6), 881–896. <https://doi.org/10.1177/1090198106297855>
- Vlassenroot, K., & Huggins, C. (2005). Land, Migration and Conflict in eastern DRC. In *From the ground up: land rights, conflict and peace in sub-Saharan Africa* (pp. 115–195).
- Walsh, G., Gwinner, K. P., & Swanson, S. R. (2004). What makes mavens tick? Exploring the motives of market mavens' initiation of information diffusion. *Journal of Consumer Marketing*, 21(2), 109–122. <https://doi.org/10.1108/07363760410525678>
- Woomer, P., Huising, J., Giller, K. E., & Et, A. (2014). *N2Africa : Final Report of the First Phase 2009-2013*. Retrieved from www.N2Africa.org
- WORLD, B. (2008). *Agriculture for Development*.
- WORLD, B. (2015). World Bank Development Indicators. Retrieved from data.worldbank.org/indicator/NY.GDP.PCAP.CD
- Zingore, S. (2011). Maize productivity and response to fertilizer use as affected by soil fertility variability manure application and cropping system in Africa. *Better Crops*, 95(1), 9.

Appendix

A1 Regional variation

The agro-climatic, geographical terrain, and institutional variation of the South Kivu province is most easily grouped into three distinct axes. The first axis stretches north of the regional capital, Bukavu, along the shoreline of Lake Kivu. The second stretches west from Bukavu into an area characterized by highlands. The third is directly south of Bukavu and consists primarily of the Ruzizi Plain which is the floor of the Western Rift Valley. Table A3 outlines the breakdown of our sample into each of the axes.

Table A1a Sample by Axe

Axe	Territory	vill_id	hh_id
North	Kalehe	16	112
Central	Mwenga	23	203
South	Uvira	52	436
Total		91	751

When looking at the primary production inputs and crop production frequencies by region, the variation is apparent between axes in Table A4. Overall the regional differences are quite strong, making it imperative that geographical fixed effects are controlled for in the analysis approach.

Table A1b Farm Characteristics and Crop Frequencies by Axe

Axe	Cassava	Soybean	Common Bean	Farm Size (Ha)	Farm Owner	Soil Quality	Hired Labor (USD)	Mean Mem. Labor	HH
North	58%	18%	83%	0.46	0.82	3.23	26.65		3.30
	(0.50)	(0.39)	(0.38)	(0.59)	(0.36)	(1.02)	(343.84)		(2.33)
West	75%	2%	64%	0.45	0.87	2.74	0.48		2.96
	(0.43)	(0.14)	(0.48)	(0.80)	(0.31)	(0.98)	(1.59)		(2.00)
South	76%	1%	39%	0.41	0.75	3.65	16.16		2.66
	(0.43)	(0.11)	(0.49)	(0.63)	(0.42)	(0.84)	(267.73)		(1.70)

Stan. Dev. in parentheses

A2 Collaborating Partners

Table A2 Collaborating Partners in the N2Africa Project of Eastern DRC

International	Local	
Wageningen UR	Catholic University of Bukavu (UCB)	IPLCI
International Center for Tropical Agriculture (CIAT)	DIOBASS	CDC
the International Institute for Tropical Agriculture (IITA)	PAD	Women for Women
the Consortium for Improving Agriculture-based Livelihoods in Central Africa (CIALCA)	SARCAF	

A3 Package Composition

Table A3 Package Composition by Partner for Extension + Subsidy Treatment Arm

Partner	Package 1	Package 2	Package 3	Package 4	Package 5	Package 6	Soybean Options	Comm. Bean Options
PAD	4 kg Soy 10 gr Inoculum 3 kg Maize SW303 5 kg NPK	4 kg Soy 10 gr Inoculum 160 lm Cassava 5 kg NPK	4 kg Comm. Bean 3 kg Maize SW303 5 kg NPK	4 kg Comm. Bean 160 lm Cassava 5 kg NPK	4 kg Soy 10 gr Inoculum 5 kg NPK	4 kg Comm. Bean 5 kg NPK	SB24 PK6	CODML (black) RWR10 (red)
IPLC / WfW	4 kg Soy SB24 10 gr Inoculum 3 kg Maize 12,5 kg NPK	4 kg Soy SB24 10 gr Inoculum 160 lm Cassava 12,5 kg NPK	4 kg com. bean 3 kg Maize 12,5 kg NPK	4 kg Comm. Bean 160 lm Cassava 12,5 kg NPK	4 kg Soy SB24 10 gr Inoculum 12,5 kg NPK	4 kg Comm. Bean 12,5 kg NPK		HM21 Marungi CODML (black)
Sarcac / Diobass	2 kg Soy 10 gr Inoculum 1,5 kg Maize SW303 6 kg NPK	2 kg Soy 10 gr Inoculum 75 lm Cassava 6 kg NPK	2 kg Comm. Bean 1,5 kg Maize SW303 6 kg NPK	2 kg Comm. Bean 75 lm Cassava 6 kg NPK	2 kg Soy 10 gr Inoculum 6 kg NPK	2 kg Comm. Bean 6 kg NPK	SB24 PK6	CODML (black) RWR10 (red)

Partner	Package 1	Package 2	Package 3	Package 4	Package 5	Package 6	Soybean Options	Comm. Bean Options
Kasheke / Sarcaf	2 kg Soy 10 gr Inoculum 1,5 kg Maize 6 kg NPK	2 kg Soy 10 gr Inoculum 75 lm Cassava 6 kg NPK	2 kg Comm. Bean 1,5 kg Maize 6 kg NPK	2 kg Comm. Bean 75 lm Cassava 6 kg NPK	2 kg Soy 10 gr Inoculum 6 kg NPK	2 kg Comm. Bean 6 kg NPK	SB24 PK6	CODML (black) HM 21

Units: lm = Linear Meters ; gr = gram ; kg = kilogram

Crops: Comm. Bean = Common Bean ; Soy = Soybean; Cassava Variety = Sawa Sawa ; Maize Variety = Ekayel de Kenya

NPK = In-organic fertilizer containing macro-nutrients of Nitrogen (N), Phosphorus (P), and Potassium (K)

Table A4 Variable names, definitions and details on variable construction or estimation

Variable	Definition	Construction Details
T	Time	t = 0 as baseline and t = 1 as endline
Primary Outcomes		
Tech Eff (USD)	Technical Efficiency in US dollar (USD) (%)	Estimated using SFA with USD value of plot as dependent variable. Plot level estimates averaged to farm level for all empirical analysis.
Tech Eff (kJ)	Technical efficiency in caloric content (kJ) (%)	See above except with caloric content as value measure
Yield (USD)	Total USD value of all farm crop output	Calculated using crop price means across both time periods. Output values normalized to plot size (i.e. yield value of kilograms per hectare)
Yield (kJ)	Total kJ value of all farm crop output	Calculated using crop specific caloric content as given in Leung (1968).
Secondary Outcomes		
Intercropping	Proportion of plots for which intercropping changed	Change variable calculated as the number of times an intercropping combination in the baseline is not observed repeated in the endline, normalized by total number of plots cultivated.
Input Use	Input Use Index Score	Calculated using Principal Component Analysis on the production function variables used to estimate technical efficiency in the Stochastic Frontier model
Bean Yield (Kg/Ha)	Total farm yield for common bean (Kg/Ha)	Calculated only for farmers who sowed common bean seeds. Normalized to yield estimates.
Fertilizer (=1)	Indicator variable for use of in-organic fertilizer on the farm	
Household		
Social Capital	Social Capital Index Score	Calculated based on responses of available connections to provide assistance in hypothetical situations of agricultural negative shocks. Calculations include number of connections available for assistance listed, traveling distance to them, closeness of relationship.
Used Credit (=1)	Indicator of use of credit in the past 12 months	
HH Head Age	Age in years of household head	Some observations listed large differences in ages between baseline and endline despite no change in responder, these outliers are corrected for by adding “2” to the baseline response.

Variable	Definition	Construction Details
HH Head Edu	Categorical variable of highest level of education attained	Responses of “None”, “Primary”, “Secondary”, and “Tertiary”. Some completion included as having attained a level. Tertiary includes undergraduate, graduate, and professional certificates.
Female (=1)	Household head is female	
Migrant (=1)	Household head is a migrant	Migrant defined as not being born or married into the village.
HH Tin Roof (=1)	Indicator for the house having a tin roof	
Agr. Coop. (=1)	Indicator for the household head being a member of an agricultural cooperative	
Production		
Crop Bundle	Total number of different crops produced on farm	
Farm Size	Total farm size (Ha)	Calculated as the sum of all cultivated plots
HH Labor	Number of household members who work on the farm	
Hired Labor (=1)	Indicator for hired labor use	
% Farm Owned	Proportion of total cultivated land owned by the household	Mean of all plots weighted by plot size
Soil Fertility	Soil fertility level on a Likert scale (1 to 5)	1 = Very infertile; 2 = Infertile ; 3 = Average ; 4 = Fertile ; 5 = Very Fertile Each plot rated and weighted mean calculated by plot size
Organic Fertilizer (=1)	Indicator for use of organic fertilizer (manure) on the farm	
In-organic Fertilizer (=1)	Indicator for use of in-organic fertilizer (NPK) on the farm	
IC	Intercropping indicators	Set of indicators for each intercropping combination observed in sample. Equals 1 if that combination was observed on any of the farm plots in a given time period.

A5 Attrition and Replacement of Households in Sample

Logit model conducted on both households departing from sample in baseline and households added to sample in endline. Standard errors clustered at the village level. No observable correlation between households dropping or adding to sample and treatment assignment.

Table A5 Correlation of Attrition and Replacement Households in Sample

VARIABLES	(1) Attrition HH	(2) Replacement HH
Extension Only	0.143 (0.275)	0.558 (0.414)
Ext + Subsidy	0.050 (0.252)	0.305 (0.409)
Constant	-1.428*** (0.206)	-2.833*** (0.314)
Observations	895	770

(1) run on baseline data only ; (2) run on endline data only

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

CHAPTER 3

Social Networks and Technology Diffusion

Evidence from a Field Experiment in the Congo

Abstract

The adoption of new technologies plays an important role in political and economic development. Social networks likely play a role in this process, though little is known about the precise role the networks play or the consequences of selecting initial recipients from different positions within the networks. We collect information about full social networks from 2563 community members in forty communities in the Democratic Republic of Congo, and implement an intervention in which ambassadors are asked to spread fertilizer. We randomize whether these ambassadors are central or isolate, allowing us to estimate the effect of network position of initial recipients. In total, 225 transfers of fertilizer packs took place. We find no evidence for an effect on fertilizer use, fertilizer knowledge, willingness to pay for fertilizer among populations, nor the spread of distribution. We do find evidence that centrality affects the speed of distribution, and the identity of the ultimate recipients of new technology or knowledge: isolated farmers tend to give to the central, but central farmers do not give to the isolated. We also document large effect-attenuation, which suggests the importance of selecting a sufficiently large set of lead community members for the spread of new technology.

Publication Status: Ross, M., P. Hofman, J. Larson, P. Van der Windt, and M. Voors. (2017) “Social Networks and Technology Diffusion: evidence from a field experiment in the Congo.” Working Paper.

1 Introduction

Economic development hinges on the spread of new information and technologies. Access to information communication technologies have been found to improve political accountability (Grossman, Humphreys, & Sacramone-Lutz, 2014) and reduce corruption (Bailard, 2010); cellphones and monitoring applications have positive effects on improving health and education outcomes by reducing absenteeism (Cilliers et al., 2014); and new agricultural technologies can boost crop output and efficiency (Pypers et al. , 2011). Social network are likely important for the adoption of this new information and technologies (Foster & Rosenzweig, 2010). Furthermore, the network position of the initial recipients have been found to be important (S. Aral & Walker, 2012; D. M. Centola & Macy, 2007). This study aims to contribute to the growing literature of empirical studies testing network entry points on information and technology diffusion.

Our study focuses on the diffusion of information and technologies within the context of agricultural development. The use of interventions that rely on social learning mechanisms, typically termed farmer-led extension, has received substantial investment in the last decades (Feder, Anderson, Birner, & Deininger, 2010; Kondylis et al., 2014). Considerable attention has been paid to the effectiveness of such programs on increasing agricultural productivity and strengthening farmer livelihoods (Cunguara & Moder, 2011; Kondylis et al., 2014; Lambrecht, Vanlauwe, & Maertens, 2016; Taye, 2013). Extension programs implicitly assume that information and innovations diffuse through community networks and interpersonal relationships. This study aims to test these underlying assumptions on the diffusion of agricultural innovations.

We explore the diffusion of in-organic fertilizer and related information on its use, proper application, costs, and benefits through community social networks in the Democratic Republic of Congo using a framed field experiment. Specifically it explores how resulting selection of network contact points based on eigenvector centrality scores varies resulting diffusion outcomes. We explore the impact of network entry through most central versus least central individuals on three different outcomes. Specifically, we analyze outcomes of community rates of fertilizer use, mean scores of fertilizer knowledge, and change in willingness to pay for fertilizer. We also explore speed and spread of distribution. Second, we explore whether attenuation of fertilizer and information occurs during the diffusion process, and whether this is different by type of

ambassador. Third, we investigate which characteristics are important to receive fertilizer and information about fertilizer.

Our findings reveal a number of new insights about the role of networks in technology diffusion. First, we show that network centrality measures are highly correlated with observable characteristics of community residents. Central farmers are more likely to hold a formal position of status in the community, are more likely to interact with community leaders, are less likely to be a migrant, are in general richer, more literate, more likely to be male, and have larger households. This corresponds closely with anecdotal work from the region. These strong correlations with observable characteristics are important in guiding implementers in selection of central individuals when network surveys are not feasible.

We also find that initial recipients with low centrality spread the chemical fertilizer faster compared to initial recipients with high centrality scores. However, we find no difference in the resulting community-level fertilizer adoption rate, diffusion width, fertilizer knowledge, willingness to pay between network-central versus network-isolate entry points. We find strong evidence of attenuation regardless of network entry point. Meaning, that the intensity of adoption, fertilizer knowledge, and willingness-to-pay outcomes weaken significantly with each step away from the initial network entry point. This suggests that the strongest social influence is stemming from the legitimization of contact with external extension professionals, which fades as information and innovation spread through the network.

Finally, we find compositional effects of diffusion. Specifically, that network centrality affects *which* community members receive the new technology and its associated information. Both centrals and isolates as entry-points prefer to share with individuals that live in close geographical proximity, are family, are field neighbors or are agricultural discussion partners. This result highlights that diffusion follows existing interpersonal relationships. However, a key outcome is that both centrals and isolates prefer to pass their information and fertilizer resources on to highly central individuals. This means that regardless of entry-point, innovations are going to cluster in the center of networks. This pooling of information and resources results in network central individuals being significantly more likely to benefit from any generated returns while periphery households are more likely to be excluded.

These results contribute to a number of literatures. First and foremost our study adds to the growing literature that explore the role of social networks on the spread of innovations within developing countries (Beaman, Benyishay, William, & Mobarak, 2015; Emerick, Janvry, Sadoulet, & Dar, 2016; Kimbrough, Smith, & Wilson, 2008; Nourani, 2016). Ancillary results explore questions of how social networks relate to individual characteristics and can be used by development practitioners in targeting and identifying key beneficiaries (Chami et al., 2014). Each of these research areas offers valuable insights in how development programs can more effectively harness social network dynamics for strengthening program impacts.

The rest of the paper is organized as follows. The next section introduces related literature and our contribution. Section 3 anchors our study in the Congolese context. Section 4 presents our research design. Descriptive statistics and results are presented in Sections **Error! Reference source not found.** and 6. We conclude in Section 8.

2 The role of social networks

There is a growing experimental literature exploring the way that social networks, and an individual's position within a network, matter for development outcomes.

2.1 The role of the initial receivers in distribution

Selection of network entry-points is believed to be a key design component for the speed at which innovations diffuse through a network. Innovation diffusion depends on a critical mass-adoption threshold being reached in order for adoption to become self-sustaining (Rogers, 1995). With smart targeting of entry-points the lag time between introduction and mass-adoption is reduced, thereby accelerating innovation diffusion (Foster & Rosenzweig, 2010). The selection of adoption-influential individuals has its origins within the opinion leadership literature (T. W. Valente & Pumpuang, 2007; Thomas W. Valente & Davis, 1999; Walsh, Gwinner, & Swanson, 2004) which has evolved from socioeconomic foundations to network theory foundations. Centola (2010) found that the probability of adoption was higher with greater exposure, leading to an argument that clustered selection of network-entry points was more effective for complex technologies. Banerjee et al. (2013) devised a new measure of centrality for determining “diffusion central”

individuals as a means of identifying network entry-points that were likely to maximize the spread of new innovations. Beaman et al. (2015) used complex-contagion simulations in identifying network entry-points and found resulting diffusion to be greater than *status quo* selection processes that do not utilize network data. However not all studies find network-based selection to be better as Aral, Muchnik, & Sundararajan (2013) find little variation in diffusion rates based on different entry-point selection criteria.

The most common design in evaluating entry-point selection methods consists of an intervention (often information) being offered to a random subset of individuals with pre- and post-intervention surveys used to measure individual exposure to the intervention, changes in knowledge or beliefs as a result of the intervention, and social ties within the sample community. Exposure is often inferred only from social closeness to the initial recipients of the intervention, and outcomes assume that the information traveled along existing links within the underlying social network, though this assumption is typically neither directly measured nor tested.

Several studies have documented the existence of peer effects using adaptations of the above design across a variety of contexts. In Cai et al., 2015, the intervention is a briefing on an agricultural insurance scheme in rural China. Their results indicated that respondents with a large number of friends who were exposed to the intensive training session are more likely to take up the offered insurance. In Oster & Thornton, 2012, the intervention is a distribution of menstrual cups to randomly selected girls across four schools in Nepal in which they found that girls with many friends who have already adopted the menstrual cup have a higher probability of also successfully adopting the menstrual cup. In Bandiera & Rasul, 2006, the probability of adopting sunflower cultivation in Mozambique is found to be a relationship shaped as an inverted-U. Initially, the probability of adopting sunflower cultivation is increasing with the number of family or friends who have already adopted, however at a certain threshold the relationship changes direction and additional connections having adopted cultivation lowers the probability of a given individual also adopting.

Banerjee et al. (2013) study the diffusion of participation in a microfinance program in 43 communities in India by inviting several key respondents to an informational meeting on the microfinance program. These few individuals were then asked to pass on the program information to the remainder of the community. Based on the observed diffusion, a predictive capturing an individual's likelihood for generating optimal

diffusion patterns was developed. They dub this measure diffusion centrality. Their model was very effective in predicting the diffusion pattern but computationally demanding, motivating their development of a simpler analogue termed Diffusion Centrality. This measure of Diffusion Centrality has a close relationship with eigenvector centrality, and under model assumptions of infinite diffusion time periods is equivalent to eigenvector centrality. This study's design holds five main advantages over existing network studies.

First, our intervention traces both the distribution of information and a physical technology, fertilizer packages. This is in contrast to Beaman et al. (2015), Kim et al., (2015) and Emerick (2016) whose outcomes are based solely on adoption rates. This is an important contribution as adoption of a technology relies on a potential-adopter shifting their belief set such that the returns of the innovation are expected to be greater relative to the *status quo*. Diffusion of innovation-related information is therefore an early condition of innovation adoption. We are able to track the distribution of both the information and the physical resource, comparing accordingly.

Second, we look at the importance of centrality. This focus is based on eigenvector centrality being equated to measures of social influence (Borgatti, 2005), making it particularly relevant as more influential individuals are thought to have a positive effect on the adoption decisions of peers. Furthermore, our design is especially favorable to detecting the importance of network-based selection of initial recipients on resulting diffusion patterns. By selecting household heads as entry points based on their having very high or very low eigenvector centrality scores, we overcome the identification challenges faced by Banerjee et al. (2013) and Cai et al. (2015). In addition, by focusing on the two extremes of the centrality spectrum we maximize the discernible differences in diffusion attributable to eigenvector centrality-based selection of entry points.

Third, we collect full household network data. There are potential biases that can result from extrapolating results based on small samples capturing only portions of larger networks (see (Sinan Aral et al., 2009; Borgatti, Carley, & Krackhardt, 2006; Chandrasekhar & Lewis, 2010)). In contrast, our data set provides a high-resolution snapshot of pre-existing community networks allowing for a detailed measure of network positions and overarching structures through which diffusion can be tracked.

Fourth, we measure social networks in the pre-intervention survey and use this information to select our farmer entry points which allows for a more causal relationship to be established. To our knowledge, a

similar network-based selection approach has only been undertaken by two prior studies, both of which analyzed alternative network measures for their selection processes. Kim et al. (2015) collected network data in 20 communities in Honduras and assign communities to one of three treatments: 1. Random Assignment 2. Community members with the highest number of social network connections (indegree) 3. Friends of randomly selected community members. They find that using indegree does not improve diffusion outcomes over random assignment, and friends of friends only increase adoption for one of their two interventions provided. Beaman et al. (2015) collected full network data in 200 communities in Malawi based on three relationships of interest: consultation on agricultural decisions, food sharing, and friendship to study an intervention aiming to increase adoption of pit planting methods. Using threshold models of contagion they simulated both simple and complex contagion models in selecting theoretically-optimal pairs of initial seeds. They find that all initial recipients selected from either of the threshold model simulations (simple or complex) outperformed the status quo method of selection by consultation with community leaders.

Finally, our design allows us to collect a wide set of outcome measures. We measure outcomes in terms of not only the magnitude of community-level diffusion and uptake, but also the composition of recipients by exploring who within the network ultimately receives either the fertilizer or the information. It is important to learn more about the types of people that are benefiting from diffusion-based interventions given concerns that resources are captured by community elites or diverted along personal channels (Olken, 2006; Reinikka & Svensson, 2005). Reaching marginalized populations or particular subsets of community members may be of particular importance to development agencies. Our design allows us to consider *which* community members gain access to and knowledge about new technologies when different network-based entry points are selected for resource distribution. Studies that aggregate community-level effects overlook this potentially crucial consequence of technology diffusion schemes.

3 Research context: Eastern Congo

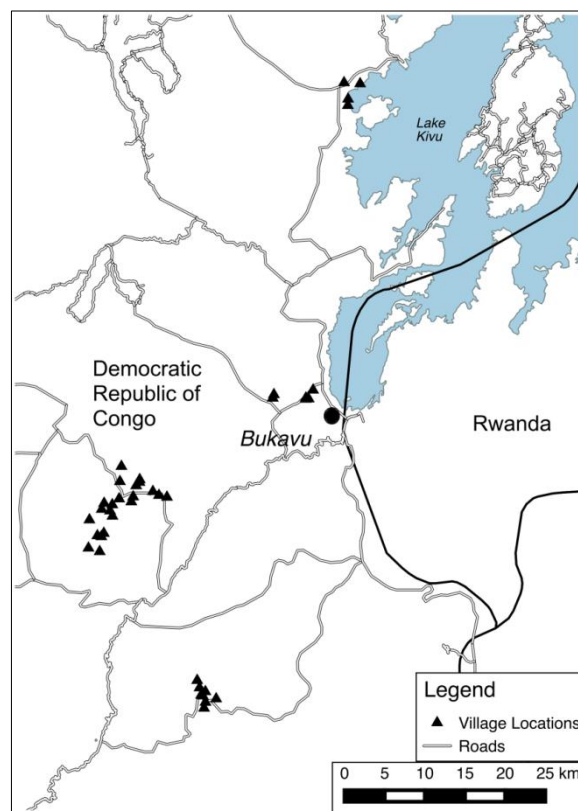
We implement our study in forty rural communities in the Congolese province of South Kivu. The location of the communities and the major regional city (Bukavu) are shown in Error! Reference source not found..¹³ Congo ranks at the very bottom in the human development index (UNDP, 2013). GDP per capita is as low as \$442 and over 85 percent of the population currently lives below the poverty line (WORLD BANK, 2015). The region has been embroiled in violent conflict, spiking during the First and Second Congolese Wars (1996-1997 and 1998-2003). The latter has been the deadliest war in modern African history (Coghlan et al., 2007). Hostilities remain up to this day. For the majority of the rural population agriculture is the main source of income and nutrition. Farms are often small and fragmented. Farms typically intercrop common beans, banana, cassava, sweet potatoes, maize and sorghum (Bulte et al. 2015). Due to the conflict, development in farming has been stagnant. Yields are low, causing widespread malnutrition. Incomes from cash crops are minimal and most farmers struggle to improve their livelihoods. Most farmers have no access to important input markets such as fertilizer and (improved) seeds (Pypers et al., 2011). As a result, the use of fertilizer is very low, with only 3% of farmers reporting having used fertilizer previously (Bulte et al., 2015). In addition to infrequent use, average quantities applied are low. On average, only 0.3 kilograms per hectare is applied, compared to 14kg per hectare for SSA as a whole and 166kg per hectare in Asia (FAO, 2015).

Communities in the South Kivu province – like in the rest of Congo – are small, typically comprised of less than 200 households (Bulte et al., 2015). Within these communities, economic and social interactions are typically local due to underdeveloped transport and ICT infrastructure. The reach of formal government is limited, and local institutions – such as chiefs - have considerable autonomy in organizing the economic and social activities of raising taxes, settling disputes and allocating communal resources (Vlassenroot & Huggins, 2005). Additionally, most agricultural activities and organizations are formed internally to a community. This combination of agricultural innovations being largely missing within the region and the

¹³ Communities were not randomly selected. Communities were selected based on the following criteria: fewer than 100 inhabitants, road access, and proximity to larger centers where research assistants could spend the night.

localized nature in which economic and social life is organized makes Congo a well-suited location to learn about how social networks shape innovation adoption. See **Figure 1** for a map of the research region.

Figure 1 Map of Research Region in Eastern Democratic Republic of Congo



4 Experimental design

As part of our field experiment, each community was visited three times. The first visit entailed a household survey with all household heads to collect social network information as well as socioeconomic and agricultural production data. The second visit occurred approximately one month later, during this visit we implemented our field experiment. The intervention consisted of training and the distribution of packets of fertilizer and information to selected farmers, which we will call our “ambassadors”. A third visit took place two weeks later.¹⁴ During this visit, all household heads were revisited and outcome data was collected.

¹⁴ For logistical reasons there is variation in how much time there was between these visits. This variation is unrelated to assignment to treatment.

Error! Reference source not found. provides an overview of the activities undertaken. We will discuss each isit in more detail now.

Table 1 Description experimental design

Visit	Day	Activity	Date
1	1	<ul style="list-style-type: none"> • Household survey • Mapping of household-level social networks • Chief survey to obtain community level information 	17 Feb – 13 Mar 2015
2	31	<ul style="list-style-type: none"> • Training ambassadors • Distribution of technology and information • Communities randomly assigned to central or isolate ambassadors • Elicit willingness to pay for fertilizer 	20 Mar – 29 Mar 2015
3	45	<ul style="list-style-type: none"> • Household survey with the same individuals as visit 1 • Track fertilizer and knowledge spillovers 	8 Apr – 29 Apr 2015

4.1 Visit 1: Measuring the social network

Our research assistants first conducted a full community census of all heads of household in the community, during which the head of household’s full name, age, and gender were recorded along with the names of any other adults in the household. Upon completion of the census, each household head was individually interviewed. The baseline survey consisted of two parts. The first part collected information about basic socio-economic characteristics, such as demography as well as income and agriculture related questions. These questions included information about fertilizer use, fertilizer knowledge and an individual’s willingness to pay for fertilizer. We will come back to these measures.

In the second part we collected social network data. Specifically, we aimed to obtain data on three types of network. The first is *family network*. Whether the head of the household is biological family with any member within the other household.¹⁵ Second, we ask about *field-neighbor network*. Whether the head of the household’s fields borders a field owned by any member of the other household. Third, we measure the *agriculture network*. Whether the head of the household discusses agricultural related topics with anybody else in the other household. We focus on these three networks because they are most closely associated to our

¹⁵ Specifically, we use whether the other person is biologically related to a maximum of the third degree (this is a well-understood term in Congo). This does not include the wife’s family; it has to be through descent.

agriculture-related intervention, which is discussed in the next section. Kendzior, Zibika, & Voors (2015) found that these three networks are the three predominant channels via which agricultural resources were shared. We conducted a number of pilots that included more networks.¹⁶ Pilot results showed that these three dimensions were the most distinct from one another and thus captured maximum variation while minimizing the number of network survey questions.

Specifically, to elicit network ties the research assistant first explained the network under study, and then moved down the complete community roster asking for each household on the list if the network applies.¹⁷ Our network data is thus at the household level for each of the three relationships covered.

4.2 Visit 2: Training and distribution of fertilizer and knowledge

The aim of this study is to understand the role of social networks in technology distribution. As a result, we undertook an intervention with a technology that is relevant to community members in Eastern Congo. We choose for the distribution of fertilizer and knowledge about fertilization to a select set of “ambassadors” per community, whom we asked to distribute this technology and information further. This intervention respond to the low yield faced by Congolese community members. In addition, the intervention mirrors closely agricultural extension programs designed to promote technology adoption that have been undertaken in the region by NGOs in order to raise yields and subsequently improving nutrition and income levels (Pypers et al., 2011). Acknowledging the promise of new technologies, development initiatives often strive to distribute them in developing countries. A typical approach first selects a subset of community members to be the initial recipients who are then tasked with spreading the technology and information about it more widely throughout the community (see Feder, Murgai, & Quizon, 2003). These programs train and provide resources to a subset of intended recipients, frequently referred to as “lead” farmers or “contact” farmers, and rely on these trained farmers to further disseminate information and resources within their community. The aim is to maximize distribution at minimal costs (Woomer et al., 2014).

¹⁶ In total we conducted three pilots. The other networks were friends (the problem was that everyone was everyone’s friend), and work on another person’s farm (that overlapped with the other networks).

¹⁷ Given the time intensive method in which we elicit network ties, we opted to measure community social networks across fewer dimensions than earlier studies (e.g. Banerjee, 2013) in efforts to minimize potential survey fatigue.

The field experiment consisted of a number of activities. In each community, enumerators gathered the three pre-selected ambassadors. These ambassadors then took part in a one hour training session, which was led by an agronomist.¹⁸ The training followed a set script (see Appendix) and the topics included information on types of fertilizer, benefits, application methods, expected market prices and access points in Bukavu. Our training sessions focused exclusively on in-organic fertilizer containing NPK for its flexibility in being applied throughout the growing season and its positive effect on yields for a range of crops grown within the region (FAO, 2000).¹⁹

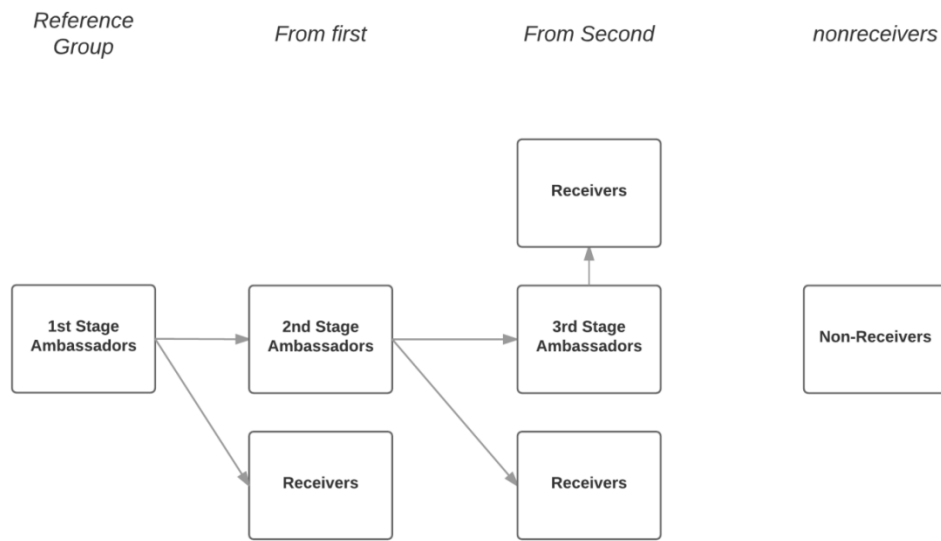
At the end of the training session, each ambassador received a single 1 kilogram bag of fertilizer that they were told was theirs to keep. Each ambassador also received three fertilizer “kits”. Each kit consisted of three 1 kilogram bags of fertilizer. All bags of fertilizer came with a sticker and a pen. We asked the ambassadors to distribute each kit to separate households in the community. Upon transfer of the kits, we asked the ambassador to record the date and time on a sticker and paste it on the kit. We also asked the ambassadors to spread the information on use of in-organic fertilizer that was provided during the training. The recipients of the kits were asked to take one bag from the kit, and distribute the two remaining bags further. These “second-stage” ambassadors were also asked to record the date and time of this transfer on a new sticker, and to distribute the knowledge of fertilizer as well. Given this structure the maximum spread is thus thirty recipients per community, or 27 transfers (nine from first-stage to second-stage ambassadors, and eighteen from second-stage to final recipients). Note that there were no sanctions or incentives imposed to ensure this structure was followed in practice.

In our analysis below, we refer to the three trained individuals as *first-stage ambassadors*. Those community members that received fertilizer and training from these first-stage ambassadors are called *second-stage ambassadors* so long as they also further distributed fertilizer. If any individual received fertilizer (from either first-stage or second-stage ambassadors) but did not further distribute fertilizer, they are referred to as *receivers*. Community members that did not receive any fertilizer are termed *non-receivers*. This structure is illustrated in **Figure 2**.

¹⁸ The agronomist was one of the enumerators who had finished a degree in agronomy.

¹⁹ NPK fertilizers are three-component fertilizers providing nitrogen (N), phosphorus (P), and potassium (K).

Figure 2 Design Structure of Fertilizer Distribution



4.3 Exogenous variation in the type of ambassador

In this study we aim to understand the importance of the position of the initial receiver in the network for the distribution of technology. A major problem research face answering this question is that the initial receiver is not likely to be random. NGOs that undertake agricultural extension programs often target specific individuals, whether they are the more literature, those chosen by the community chief, etc (Feder et al., 2010; Simpson et al., 2015). In order to be able to make causal claims, we randomly assign the initial entry point.

As a first step we collapse the three networks into a single combined network. We do so because by examining multiple networks, we look at more possible channels of interaction. Therefore, we can capture more of what is happening in the social networks in the community. By combining the three networks we examine all possible ways of interacting that we think are likely to explain agricultural distribution behavior. Specifically, for an relationship to be present in the combined network graph it must be present in at least one of the three networks. Since a relationship need not be reciprocated for a node (head of household) to transfer information or resources, we convert our network data from directed edges to undirected edges,

meaning if either node claims a connection to another node, we will treat it as two-way link between the two.

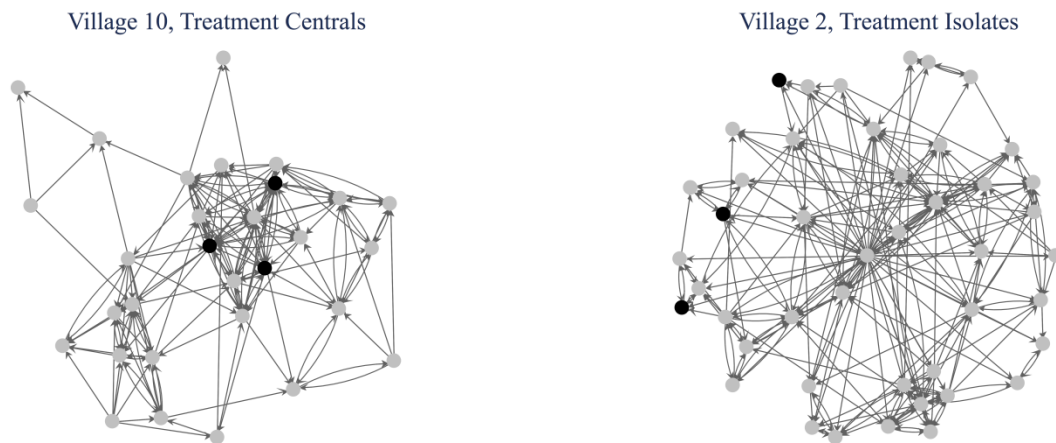
Subsequently, we assign to each individual their position within the network. In this study we are particularly interested in the centrality of an individual. We expect that the centrals are more influential (Bonacich, 2007). This should lead to higher uptake and knowledge in communities where central farmers were trained, compared to communities where isolate farmers were trained. Specifically, in this study we focus on a single measure of an individual's position within a network structure: eigenvector centrality (Bonacich, 1972). This centrality measure is based not only on the number of edges that a given node has (as in degree centrality), it also weights each edge between nodes by the degree of the node that the edge leads to. For this reason, eigenvector centrality is often used to proxy for an individual's level of influence within their social network as nodes can be said to be more influential with each other well-connected node it is linked with. In the case of technology diffusion, a highly eigenvector central node may, by nature of its access and connectedness, be a more compelling source of novel information for adoption of the relevant technology. Alternatively, nodes with low eigenvector centrality may, by nature of homophily, be better positioned to reach nodes in the network that lay on the periphery of the network structure (Aral et al., 2009).

As a last step we assign our communities to one of two treatments.²⁰ In twenty randomly selected communities we choose the three most central individuals as ambassadors – and call them “central ambassadors”. In the other twenty communities, we choose the three least central individuals as ambassador – we call them “isolate ambassadors”.²¹ Error! Reference source not found.3 illustrates a network map for two communities, one from each treatment group. The map is structured so that individuals with fewer connections are located on the periphery while those with numerous connects are located in the center of the network map.

²⁰ We also randomly assigned our five research teams to these community to deal with any worries related to enumerator effects.

²¹ In some cases, one (or more) of individuals with the highest (lowest) centrality score were not present in the community. In such case, they were replaced by the next highest (lowest) from a backup list.

Figure 3 Example Community Social Network Structures by Treatment Type (Isolate v Central)



Error! Reference source not found. presents information for a number of key respondent characteristics, outcome variables, and social network based on pre-treatment information across the two treatments. We find that the randomization was successful in creating two groups that are similar. Across nineteen measures we find differences for only one variable. In-degree, or the number of edges that lead towards a node, is higher amongst households in communities where isolates were selected as entry points. The presence of an imbalance between treatment groups, even when assignment to treatment is randomized as in our study, is to be expected when comparing data over many characteristics. However, as this is likely to be correlated with treatment we add indegree in all our analyses to improve power.

Table 2 Descriptive statistics and balance

	Treatment Centrals			Treatment Isolates			
Variable	N	Mean	SD	N	Mean	SD	P-value
<i>Individual and community level characteristics</i>							
Age	1,191	46.75	17.87	1,336	45.16	17.18	0.149
Gender (1=female)	1,191	0.32	0.47	1,336	0.32	0.47	0.858
Literate (yes=1)	1,167	0.49	0.50	1,298	0.43	0.50	0.311
Ethnicity (1=minority group)	1,195	0.08	0.28	1,338	0.18	0.39	0.307
War Exposure (max 4.5)	1,171	2.41	1.07	1,304	2.22	1.01	0.249
Household Size	1,171	6.29	3.14	1,304	6.37	3.18	0.748
Income Index	1,135	0.00	1.00	1,255	-0.03	0.91	0.648
Likely lead farmer (1=yes)	1,195	0.23	0.42	1,338	0.21	0.41	0.480
Chief (1=yes)	1,171	0.02	0.13	1,304	0.02	0.13	0.899
Migrant (1=yes)	1,191	0.35	0.48	1,335	0.31	0.46	0.181
Community Size	20	62	19	20	68	13	0.246
<i>Outcome variables</i>							
Ever used fertilizer (yes=1)	1,158	0.06	0.24	1,285	0.07	0.26	0.424
Fertilizer Knowledge Score (max 7.5)	1,166	1.27	1.47	1,299	1.26	1.44	0.954
Willingness to Pay (RCS, in US\$)	335	1.49	1.34	339	1.59	1.46	0.664
<i>Social network characteristics</i>							
Centrality	1,195	0.34	0.27	1338	0.32	0.26	0.591
In-degree	1,195	7.79	5.08	1338	10.02	7.10	0.029***
Density	20	0.14	0.07	20	0.15	0.07	0.620
Clustering coefficient	20	0.60	0.18	20	0.62	0.20	0.783

Note: P-values are from regressions with standard errors clustered at the community level (where appropriate). There are some small deviations from the PAP here, where we specified simple t-tests without clustering. Land size is restricted to maximum 100 km². Income index is a normalised index (for the control group) calculated using the user-written Stata program WMEANEFECTS and comprises: # chickens owned, # goats or sheep owned, # cows owned, land size, land access.

4.4 Visit 3: collecting outcome measures

About two weeks after the field experiment, we revisited all communities to collect outcome measures.

Specifically, we are interested in individuals' use of fertilizer, knowledge about fertilizer, and willingness to pay for fertilizer. Furthermore, we are interested in the speed and spread of the technology throughout the community. To obtain this information our enumerators undertook two activities in each community: a tracking exercise and a second household census.

The packets of fertilizer distributed during the intervention were tracked down by research assistants.²² The starting points were the three ambassadors to whom fertilizer was initially distributed. We then asked to whom they gave fertilizer and followed up with the named recipient. We asked each fertilizer recipient whether, when and from whom they received a packet of chemical fertilizer, and to whom they in-turn distributed any fertilizer. These additional recipients were then also visited and interviewed until all recipients had been identified and receipt of fertilizer confirmed. Based on this data we created two outcome variables of interest: speed of diffusion and spread of diffusion.

Speed of diffusion was measured as those giving fertilizer had been asked to note on a provided sticker, the date and time a package transfer was made. We copied this information and calculated for each packet of fertilizer the number of days that had occurred between the recorded date of receipt and the date the intervention was conducted within the community during which all fertilizer packets were initially provided.²³

To calculate the *spread of distribution* we take all nodes that received in-organic fertilizer at any point during the intervention and calculate the geodesic distance between receivers. The geodesic is the minimum number of edges that need to be followed in connecting one node with another. The geodesic is then divided by the total number of nodes in that community (to make the results comparable across communities). If there is no shortest path possible (because the receivers are in unconnected parts of the network) this variable takes the maximum value of 1.

Once all fertilizer packets had been traced, our research assistants re-surveyed all households in the community. As part of the survey, we collected data was collected on three outcome variables. The first is

²² We only tracked fertilizer packages. Not the knowledge (i.e. with whom did you talk?), that was collected using a survey.

²³ If the recorded time of transfer was after 5pm, we recorded an additional half day as having passed.

fertilizer use. We ask respondents whether they had applied chemical fertilizer on any of their fields during this agricultural season.

Next, we asked about *fertilizer knowledge*. We asked each respondent a set of questions to capture each household head's knowledge on chemical fertilizer use, benefits, and market sourcing. Specifically, the questions covered the expected effects of chemical fertilizer, timing of application, methods of application and market availability and pricing.²⁴ Points were awarded for correct responses to generate a fertilizer knowledge score ranging from 0 to 7.5.

Finally, we ask about individuals' *willingness to pay* for fertilizer. To do so we used the Randomized Card Sorting (RCS) method, as described in Smith (2006)²⁵, with eighteen respondents from each community. These eighteen respondents were chosen by sorting all household heads by eigenvector centrality scores and selecting the top six, bottom six and middle six household heads. Using ten choice cards which displayed potential fertilizer prices that increased stepwise from 100 Congolese Fc (about 0.90US\$) to 5000 Fc (about 4.50US\$), each of the eighteen selected participants were asked which of the fertilizer prices displayed they would be willing to pay for a one kilogram bag of NPK fertilizer. Cards could be sorted into one of three piles: would pay, would maybe pay and would not pay²⁶.

It is important to note that we asked the exact same household survey questions before the onset of the field experiment. As a result, we have panel data that allows us to identify changes in indicators of interest that arise due to the chemical fertilizer distribution intervention.²⁷

²⁴ We tally and weigh the correct answers to create a knowledge score. Specifically, it is based on the following questions. What is the effect of fertilizer on yields? (A: increases yields, 1 point) What are other effects of fertilizer? (A: earlier harvest, 0.5 point. A: Kills seeds, 0.5 points. Multiple possible). When is it effective to apply fertilizer? (A: before planting, 1 point. A: during planting, 1 point. A: after planting, 1 point. Multiple possible). What is the best method to apply fertilizer? (A: mix in soil before planting, 0.5 point. A: Put next to the seed/plant, 0.5 point. A: Throw on the field, 0.5 point. Multiple Possible) What is the price of fertilizer in Bukavu? (A: between 1.5 and 1.7 dollars, one point. A: between 1.2 and 1.5 dollars, 0.5 point. A: between 1.7 and 1.9 dollars, 0.5 point).

²⁵ A challenge with measuring Willingness to Pay measurements is starting point bias (or anchoring). Respondents rarely deviate far from the initial valuation posed by the researchers. By giving all the amounts to the respondent in a random order this is prevented. Smith (2006) compares three approaches that use payment cards, and finds evidence that randomizing the order of the cards yields the most valid results.

²⁶ Included in the baseline survey was a simple open-response questions asking the amount the respondent would be willing to pay for a one kilogram bag of NPK fertilizer. The average response for this question was \$1.00, significantly lower than the price elicited using the Random Card Sorting method conducted with the 18 community members.

²⁷ The survey also asked whether the household had received fertilizer in the previous weeks. As a result, we have two sources on whether the household received fertilizer: 1) from the tracking exercise, and 2) from the endline survey. In the current analyses, we use the first. There seem to have been some data entry errors providing a large mismatch between both data sets, ie when comparing the tracking data (whom did you give fertilizer to) to the endline data (who did you receive fertiliser from) do not overlap to a large degree. We have to go back to the data entry forms to check.

4.5 Attrition

We undertook great efforts to interview all heads of households. Our rule was that if the household head was not available, we would return later (often during the weekend as respondents were more likely to be present). If the head of the household remained unreachable, we looked for an adult replacement within the household. In to the first household census, in about 20% of the cases the head of household had to be replaced with a family member, generally the spouse.²⁸ Upon replacement, we asked the replacement about the characteristics (including the network characteristics) of the head of the household. In total, we were able to data from 97% heads of households, which is very high compared to similar studies.²⁹ In total, we recorded information on 2,563 individuals and 20,596 social ties.

Related to our second household survey conducted during the third visit, we record modest attrition. In total, 9% of our baseline respondents were not present. We do not find that there was selection in attrition.³⁰ We also do not find that the level of attrition is correlated with treatment status.

4.6 Manipulation check

Before moving to the results, we first explore whether the ambassadors in fact distributed the fertilizer. We find that they did. In total, our dataset records 225 transfers of fertilizer. In total, 86% first-stage ambassadors distributed fertilizer.³¹ Interviews reveal that those who did not distribute their fertilizer kept it for their own use, or said that they were waiting until the next planting period to distribute the fertilizer. We find that central ambassadors donate to an average of 3.1 other communities, and the average kilograms per recipient is 1.6 kilograms. Isolate ambassadors distribute to an average of 2.8 individuals, with an average of 2 kilograms per recipient. This last difference is significant at the 95% confidence level using a simple t-test. We find that distribution beyond the first-stage ambassador is much lower. Second-stage ambassadors

²⁸ We asked the replacement the reason for the absence of the head of the household. The most common reasons mentioned are: visit to the household's fields, visit to Bukavu, and temporary outmigration for work.

²⁹ Beaman et al (2015), for example, were only able to reach 80% of targeted respondents.

³⁰ In **Table 7** in the appendix we regress attrition on a set of baseline characteristics and find only one correlation: dropouts have had slightly lower war exposure. The difference is however small. Further, we also collected some data on participants we could not reach in any of the field visits. This data was collected from a neighbour, family, or the community chief. We then compare this group against all who participated in the endline survey. Absence is very low (3%) migrants and farmers are also less likely to be absent.

³¹ These are 97 of 116 first stage ambassadors (we lack data for 4 ambassadors)

distributed onward in only 20% of cases. When they did, they distributed on average to 1.8 people and in total 1.1 kg.

5 The Sample: Community members and ambassadors

We present information about the community members and ambassadors in our study communities.

5.1 *The Community Members*

Error! Reference source not found. presents descriptive statistics for our respondents. Respondents are on average 46 years old, predominantly illiterate with only 32% of the household heads being female and roughly 33% of the household heads being migrants into the community. Households typically comprise of six members (including all wives and children). Most respondents have been exposed to conflict in a number of ways including being forced to flee, losing a family member or sustaining an injury. We reduced exposure to conflict to a single metric by constructing a composite index of the weighted sum of all experienced incidents of conflict.³² On average, sample members reported having experienced at least one such conflict-related event. As the average community size within our sample is 65 households, by construction roughly two percent of the sample hold the position of community chief. The average respondent's farm size is five square km.³³ Following Banerjee et al. (2013), in order to assess whether someone was viewed as influential among farmers, we asked *If we wanted to spread information about a new agricultural technique, to whom should we speak?* We did not restrict nomination to household heads but recorded any name provided by each respondent. In total 23% of the household heads were mentioned at least once as being the best candidate for spreading agriculturally-relevant information.

Turning to baseline values of our main diffusion outcome variables. Error! Reference source not found. shows that only about 6% of households had previously applied fertilizer to their fields. Overall, fertilizer knowledge in the baseline period was low with an average sample score of 1.27. Finally, based on the

³² Saw Fighting, 0.5 point. Property Damage/Loss, 0.5 points. Injured in the war, 1 point. Family member injured in the war, 0.5 point. Family member killed in the war, 1 point. Migrated because of the war, 1 point.

³³ We remove several outliers of farms over 100 km². In our analyses that require land size as a predictor we will use the measure without outliers. We did not account for this in the pre-analysis plan.

randomized card sorting method, the average willingness to pay reported was approximately \$1.50 for 1 kg, roughly the equivalent of the market price of fertilizer.³⁴

Finally, these community members live in closely knit communities. Density is the percentage of possible edges that are realized and averages 15% for our sample communities. The clustering coefficient is a measure of the proportion of possible triads being present within each community network. Our sample average for the clustering coefficient is high at 61%, indicating that there is a 61% probability that if household A is connected to both household B and household C, that household B and C are also connected. The average community member has a centrality of 0.33, and an indegree of 8.9.

5.2 *The ambassadors*

We briefly discuss our (first-stage) ambassadors. **Table 3** presents information about central and isolate ambassadors based on information collected before the intervention. As expected, we find that central and isolate ambassadors differ significantly on a considerable number of personal characteristics. Central ambassadors are more likely to be named as leaders among farmers, male, literate, richer and head larger households. Also, compared to isolate ambassadors, centrals are more likely to hold formal positions of authority, including being community chief, have more frequent interactions with the community chief, and are more knowledgeable about chemical fertilizer.

Isolate ambassadors are more likely to be migrants. Only 11% of central ambassadors are migrant, while this increases to a full 49% for isolate ambassadors, a difference that is statistically significant at the 99% significance level. This finding corresponds closely to previous work on Congo that highlight that migration status are important determinants of social and economic life, and suggests that these cleavages permeate even into the position an individual holds within their larger community social network.

Finally, we verify that our ambassador selection process was successfully by comparing network position of our ambassadors. **Table 3** shows that both the in-degree and the centrality of the central ambassadors are significantly higher (0.88) than isolate ambassadors (0.10).³⁵

³⁴ The market price in Bukavu for a kilogram of fertilizer was \$1.70 or 1 kg at the time.

³⁵ We have one outlier, with one supposed isolate ambassador having almost the maximum centrality score. It is possible that the research assistants trained the wrong person to be an ambassador. Incorporating this outlier in the analysis gives conservative estimates about treatment effect.

Table 3 Ambassador characteristics

	Centrals			Isolates			
Variable	N	Mean	SD	N	Mean	SD	P-value
<i>Individual characteristics</i>							
Age	56	50.2	13.8	57	46.1	18.4	0.292
Gender (female=1)	56	0.13	0.33	57	0.51	0.50	0.000***
Literate (yes=1)	56	0.54	0.50	57	0.32	0.47	0.060*
Ethnicity (1= minority group)	56	0.05	0.23	57	0.14	0.35	0.342
War Exposure (max 4.5)	56	2.63	0.97	57	2.26	0.89	0.098*
Household Size	56	7.39	3.97	57	5.58	3.01	0.019**
Income Index	53	0.28	0.86	56	-0.27	0.80	0.002**
Likely lead farmer (yes=1)	56	0.52	0.50	57	0.14	0.35	0.000***
Community Chief (yes=1)	56	0.18	0.39	57	0.00	0.00	0.000***
# times talked to chief in last month	46	11.3	10.5	57	6.0	8.7	0.020**
Migrant (yes=1)	56	0.11	0.31	57	0.49	0.50	0.000***
Indegree	56	14.3	7.0	57	4.5	4.0	0.000***
Centrality	56	0.88	0.11	57	0.10	0.14	0.000***
<i>Outcome variables</i>							
Used fertilizer (yes=1)	56	0.09	0.29	57	0.04	0.19	0.420
Fertilizer Knowledge (max 7.5)	56	1.71	1.42	57	0.80	1.19	0.001***
Willingness to Pay for fertilizer (US\$)	49	1.80	1.69	55	1.46	1.59	0.447

Note: P-values are from regressions with standard errors clustered at the community level. War exposure is based on one question where respondents could give multiple answers. Each answer has a score attached, and this variable sums those scores together. Saw Fighting, 0.5 point. Property Damage/Loss, 0.5 points. Injured in the war, 1 point. Family member injured in the war, 0.5 point. Family member killed in the war, 1 point. Migrated because of the war, 1 point. Income index is a normalised index (for the control group) calculated using the user-written Stata program WMEANEEFFECTS and comprises: # chickens owned, # goats or sheep owned, # cows owned, land size, land access.

6 Empirical Models

We first examine the effect of network position on technology diffusion. In exploring this relationship we begin by analysing equation (1) where Y_{ijt} is a vector of the three outcomes of fertilizer adoption, fertilizer knowledge, and willingness-to-pay for fertilizer for individual i (with $i=\{1,\dots,n\}$) in community j (where $j =$

$\{1, \dots, 40\}$ at time t (where $t = 0$ for baseline and 1 for endline). As our assignment to treatment is randomized, a causal relationship can be measured using a difference-in-difference model for estimating the treatment effect.

$$Y_{ijt} = \beta_0 + \beta_1 T_j + \beta_2 \text{Post}_t + \beta_3 \text{Post}_t * T_j + \partial_k + \varepsilon_{ij} \quad (1)$$

T_j is our treatment variable at the community level, which takes the value $T_j=1$ if the first-stage ambassador is an isolate, and $T_j=0$ if the first-stage ambassador is a central. Post_t is a dummy that signifies the post intervention (endline) period. The variable of interest is the interaction between treatment and a post-intervention dummy ($\text{Post}_t * T_j$). We expect the coefficient on this term, β_3 , which represents the treatment effect, to be negative. This is due to our hypothesis that central ambassadors are more effective at increasing adoption and communicating knowledge (Borgatti, 2005). ∂_k are dummies capturing research team fixed effects, as it is the variable on which our random assignment to treatment was blocked. ε_{ij} are Newey-West standard errors robust to heteroskedasticity and autocorrelation, clustered at the community level. We estimate with individual fixed effects to correct for idiosyncratic differences³⁶.

In exploring relationships of network entry point on outcomes of diffusion speed and wideness of distribution, we estimate Equation (2).

$$Y_l = \beta_0 + \beta_1 T_j + X_{ij} + \partial_k + \varepsilon_{ij} \quad (2)$$

, in which Y_l is a vector of outcomes capturing speed of diffusion and wideness of distribution. As we do not have this data over time, our data does not lend itself to a difference-in-difference approach. However as our data is randomized, we can identify causal relationships with an OLS regression on endline observations. Within Equation (2), T_j is our treatment indicators assigned for community j , while X_{ij} is respondent i 's indegree within community j . Indegree is included as a regression covariate as it was found as not being orthogonal to treatment within our balance analysis (Table 2).

To examine the effect on intervention attenuation we examine models 3 and 4.

³⁶ Random effects gives qualitatively similar results.

$$Y_{ij} = \beta_0 + \beta_1 \mathbf{X}_i + \varepsilon_{ij} \quad (3)$$

where, Y_{ij} captures the same three outcomes as seen in Equation (1). \mathbf{X}_i is a vector capturing our four types of recipient groups: (i) those that were trained by our agronomists (first-stage ambassadors), (ii) those that received from first-stage ambassadors (second stage ambassadors), (iii) those that received from second stage ambassadors, and (iv) non-receivers (See Figure 3). We expect the coefficient on each coefficient in the vector of estimates, β_1 , to be increasingly negative. This is based on a hypothesis that the strength of the intervention will be strongest for those directly receiving it with decreasing impact the further it moves from the network contact point. ε_{ij} are standard errors clustered at the community level.

While Equation (3) analyses overall attenuation, we turn our attention to whether attenuation levels vary by centrality level of network contact points. Specifically we estimate

$$Y_{ij} = \beta_0 + \beta_1 X_i + \beta_2 X_i * T_j + \beta_3 T_j + \theta_k + \varepsilon_{ij} \quad (4)$$

where, Y_{ij} is the same vector of outcomes as Equation (1) and (3). Added to Equation (4) are interaction terms between treatment assignment and each of the three receiver-type dummies. We expect coefficient estimates within vector β_2 to be negative. θ_k captures research team fixed effects and ε_{ij} captures standard errors clustered at the community level

Finally, we conduct an exploratory analysis on what factors are relevant for who ambassadors decided to transmit information and resources to. We initially estimate Equation (5) which does not distinguish between ambassador centrality type. We conduct this analysis by exploring each potential pairing of each ambassador with every other household head within the community on outcomes of whether fertilizer was transmitted or information transmitted between the pairing.

$$Y_{aij} = \beta_0 + \beta_1 X_{aij} + V_j + \varepsilon_{aij} \quad (5)$$

where, Y_{aij} is an indicator taking value “1” if fertilizer (information) was transmitted between ambassador a (where $a = \{1, 2, 3\}$) to household i (with $i = \{1, \dots, n \mid i \neq n\}$) in community j (with $j = \{1, \dots, 40\}$). X_{aij} is a row vector of a set of variables we expect to predict ambassador giving behavior. These include whether a and i are linked as family, field neighbors, agricultural discussion partners within their social network, or

members of the same agricultural group as well as the centrality score of individual i and individual i 's age, gender, wealth index score, leadership status, migrant status, and ethnic minority status. V_j is a vector of community dummies to correct for village-specific gifting norms. We cluster errors at both ambassador and receiver level³⁷.

To examine whether central and isolate ambassadors had different deliberations regarding who to give to, we also include interactions with treatment in equation 6. Significant positive (negative) coefficients in β_2 will show that isolates are more (less) likely to give to villagers with this characteristic than central ambassadors.

7 Results

We present our results in three parts. The first part explores the causal impact of an ambassador's network position on the use of fertilizer, the knowledge about fertilizer, and individuals' willingness to pay for fertilizer. We also explore whether the speed and spread of distribution is different depending on the type of ambassador. Second, we look at attenuation of fertilizer use and knowledge. The final part of this section explores who receives fertilizer, and whether this differs by ambassador type.

7.1 *The impact of network position on technology diffusion*

We first explore the causal effect of an ambassador's network position. Information about fertilizer use, fertilizer knowledge, and willingness to pay for fertilizer comes from individuals. Furthermore, we asked about this both before and after the intervention and we thus estimate a difference-in-difference model.

The first three columns in Error! Reference source not found. present the results. Column 1 shows that the intervention generated a significant increase in fertilizer use, raising fertilizer adoption by 8% compared to baseline levels. This more than doubles prior fertilizer use rates (6%). Although comparison with the theoretical maximum of every kilogram of fertilizer being used by a different household in an average

³⁷ This is different from the PAP in several ways: we said we would check for differences across treatment using a Chow test. See model 6 for the updated approach. We also said we would weight for the probability to be selected based on community size. This is not possible when clustering at two levels. In addition, we dropped variables in changes due to high multicollinearity. We added whether a receiver belonged to ethnic minority in community.

community is 46% (30/65). So adoption levels are considerably lower than intervention resources would suggest. This result is statistically significant at the 99% confidence level. However, we do not find any evidence that fertilizer use was different between those communities where we trained central ambassadors compared to those communities where we trained isolate ambassadors.

Table 4 Ambassador's network position and distribution

	(1) Fertilizer Use	(2) Fertilizer Knowledge	(3) Willingness to Pay	(4) Distribution Speed	(5) Distribution Spread
Treatment*Post	0.000 (0.036)	0.021 (0.144)	0.142 (0.246)		
Post	0.084*** (0.022)	0.696*** (0.110)	-0.199 (0.175)		
Treatment (1= isolate ambassadors)				-4.196*** (1.049)	0.081 (0.075)
In-degree				-0.036 (0.058)	-0.005 (0.004)
Constant	0.066*** (0.009)	1.274*** (0.034)	1.563*** (0.054)	9.476*** (1.212)	0.252 (0.172)
Difference-in- difference	Y	Y	Y	N	N
# Clusters	40	40	40	33	NA
Observations	4677	1198	4708	220	39
Level	Indiv.	Indiv.	Indiv.	Transfer	Community
Mean of dependent variable (centrals)	0.06	1.27	1.49	7.77	0.05

* p < 0.10, ** p < 0.05, *** p < 0.01

Note: OLS regressions of ATEs of treatment, fixed effects. Research team fixed effects included. Robust standard errors in parentheses clustered at the community level. We deviate from the PAP here. In-degree was added to control for imbalance at baseline.

Column 2 shows the impact of ambassador position on fertilizer knowledge. We find that the intervention increased fertilizer knowledge by 0.7 points to a new average of 1.9 points (out of 7.5). This means that at the community level, respondents were able to name about one additional advantage of in-organic fertilizer, or knew the price more precisely. This is statistically significant at the 99% confidence level. Again, we do not find that the network position of the ambassador was associated with a significant difference for the knowledge outcome.³⁸

In column 3 we find that individuals' willingness to pay was unresponsive both to the intervention and to treatment. This could be due to opposing forces occurring simultaneous. Improved knowledge on the

advantages of fertilizer is likely to increase an individual's willingness to pay for fertilizer while increases in local supply of fertilizer will have a potentially negative effect. We are unable to decompose these effects.

Next, we look at the speed and spread of fertilizer diffusion. Information for both variables was obtained only after the intervention, and we thus estimate a simple cross-sectional model. Speed of distribution is measured at the transfers level, while spread of distribution is measured at the community level. The last two columns of Error! Reference source not found. present the results.

Column 4 shows the effect on the speed of diffusion.³⁹ We find that isolate ambassadors completed their chemical fertilizer distribution four days faster than central ambassadors. This result is statistically significant at the 99% significance level. This is opposite to what we expected. We expected that central ambassadors would distribute faster, as they had more connections and thus more potential people to whom they could distribute fertilizer.

Finally, column 5 looks at the difference in the spread of distribution. We find that isolate ambassadors appear to distribute wider than central ambassadors. In central communities the path between all receivers is 25% of the maximum length possible, while in isolate communities it is 8 percentage points higher. This implies that the treatment has diffused further in the community. It makes sense that isolates lead to further diffusion: centrals are often clustered together in the center of the network map, while isolate ambassadors can be anywhere on the fringes of the network. This result is however not statistically significant. As this variable is at the community level, our statistical power is low, and we find high standard errors for all of our estimated coefficients.⁴⁰

In summary, we find significant effects of the intervention on adoption and knowledge of chemical fertilizer. However we do not find that this differs when the initial entry point are central head of households or isolates. We do see that isolates are much faster at distributing their chemical fertilizer, and they appear to spread wider, though this last result is not significant because of low power.

³⁹ The number of observations is lower here because it was not registered in all cases. This might be because ambassadors forgot to write the time down or because they did not know what the time was.

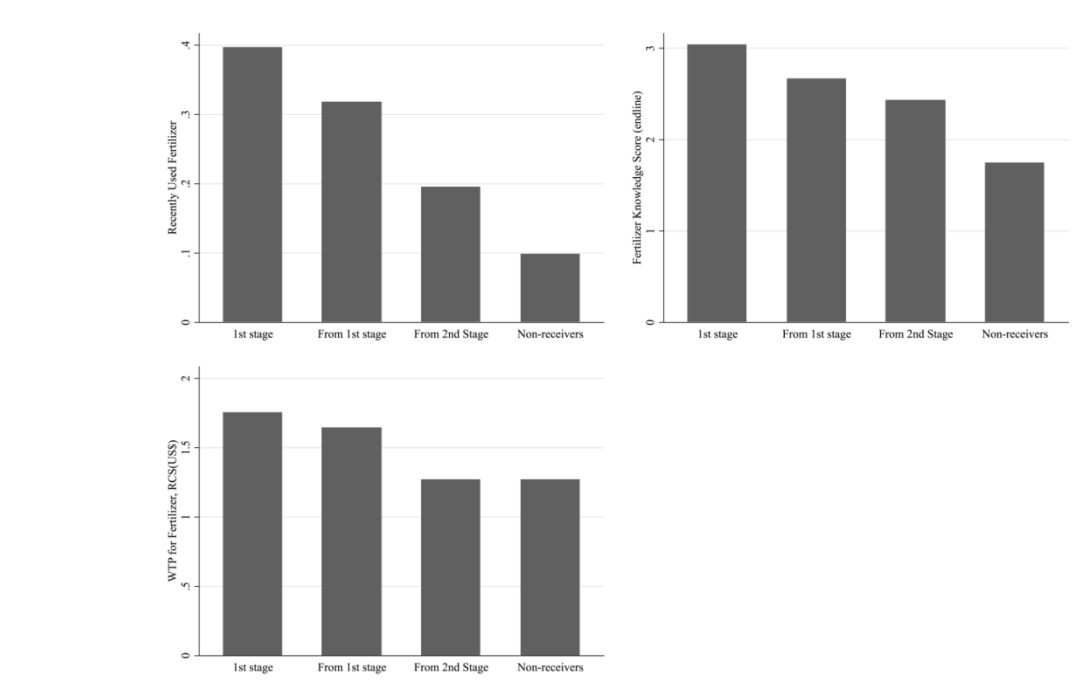
7.2 *Treatment attenuation*

The next result that we explore is attenuation of distribution. We expect fertilizer use, fertilizer knowledge and willingness to pay for fertilizer to diminish as it spreads through the community. In other words, an individual that receives information about fertilizer use directly from the agronomist is more informed than an individual that received the intervention information indirectly through network diffusion. As centrality is closely correlated with socioeconomic indicators of education, the quality of information processing and transmission behind attenuation may vary by the centrality of network contact points. We explore both now.

We divide each community into four groups: 1) first-stage ambassadors, 2) those who received from first-stage ambassadors, 3) those who received from second stage ambassadors, and 4) non-receivers⁴¹. Error! eference source not found. plots our main outcome variables – fertilizer use, fertilizer knowledge, and willingness to pay for fertilizer – across these four groups. We find that the intervention weakens when moving away from the entry point. The first-stage ambassadors have the highest level of fertilizer use, knowledge and willingness to pay. This is lower for those that received fertilizer from the second stage ambassadors, those that received from the first-stage ambassador score lower, and those that received from the second-stage ambassador score even lower. Those that are not-receivers score the lowest across all three outcome measures.

⁴¹ By design, we do not have third stage ambassadors in our data, though in practice we have one case where this happened. We lump these in with the group that received from second stage ambassadors.

Figure 6 Fertilizer Use, Fertilizer Knowledge, and Willingness-to-pay by Ambassadors, Receivers, and Non-Recipients



In columns 1 to 3 of Error! Reference source not found. we test whether these differences are statistically significant. We see outcomes taper off strongly. The coefficients are negative and increasing in size. Recipients from second stage ambassadors and non-recipients have a much smaller probability of applying in-organic fertilizer or of being equally informed on its use. Willingness to Pay follows a similar pattern, but standard errors are large. First stage ambassadors used chemical fertilizer in about 40% of the cases, while those who receive from first-stage ambassadors the adoption rate drops 8 percentage points (32%). For those receiving from the 2nd stage ambassadors the percentage is 20%. For non-receivers it is 10%. There is no significant difference between 1st stage those receiving from them. Fertilizer knowledge follows a similar pattern. 1st stage ambassadors get 3 points (out of 7.5), while non-receivers get about 1.3 points less. Willingness to pay has no clear pattern, which is unsurprising as we found no effect of our intervention on willingness to pay. In sum, we find clear and strong attenuation for adoption and knowledge. The implication of this is that an intervention will not cascade infinitely throughout a community. Having second-stage ambassadors also provide trainings will not be effective, and perhaps educating more trainers in the first stage is required to ease the teaching burden on 1st stage ambassadors.

Table 5 Centrality and intervention attenuation

	(1) Fertilizer Use	(2) Fertilizer Knowledge	(3) WTP	(4) Fertilizer Use	(5) Fertilizer Knowledge	(6) WTP
Received from 1 st stage ambassadors	-0.079 (0.061)	-0.368** (0.147)	-0.111 (0.176)	-0.101 (0.078)	-0.465** (0.190)	-0.290 (0.259)
Received from 2 nd stage ambassadors	-0.201*** (0.059)	-0.599*** (0.201)	-0.484*** (0.151)	-0.198** (0.083)	-0.843*** (0.277)	-0.377 (0.245)
Non-Receiver	-0.298*** (0.052)	-1.284*** (0.151)	-0.485** (0.210)	-0.310*** (0.077)	-1.299*** (0.193)	-0.391 (0.318)
Treatment (1= isolate ambassadors)				-0.003 (0.116)	-0.074 (0.289)	0.072 (0.287)
x Received from 1 st stage ambassadors				0.027 (0.125)	0.175 (0.315)	0.295 (0.359)
x Received from 2 nd stage ambssadors				-0.017 (0.118)	0.340 (0.385)	-0.201 (0.304)
x Non-Receivers				0.028 (0.109)	0.032 (0.306)	-0.198 (0.379)
Indegree				0.006** (0.002)	0.008 (0.006)	0.016 (0.012)
Constant	0.396*** (0.056)	3.033*** (0.140)	1.751*** (0.147)	0.282*** (0.085)	2.721*** (0.172)	1.583*** (0.252)
Observations	2295	2305	525	2236	2245	525
Number of Clusters	40	40	40	40	40	40
P	0.003	0.099	0.004			
fromfirst=fromsecond						
P	0.031	0.000	0.997			
fromsecond=nonreceiv er						

* p < 0.10, ** p < 0.05, *** p < 0.01

OLS Regressions with the three receiving groups as explanatory variables. The omitted category are the first-stage ambassadors. N for column 3 and 6 is lower because it is only for the 18 game participants.

Robust standard errors in parentheses clustered at the community level. Research team fixed effects included for columns 4-6.

Finally, we investigate whether attenuation varies by the network position of the first-stage ambassadors. Results are presented in columns 4 to 6 of Error! Reference source not found.. We find no evidence that entral ambassadors are better than isolate ambassadors at sustaining the effect of the intervention.

7.3 Predicting Diffusion

We now investigate what characteristics determine obtaining fertilizer or knowledge about using fertilizer.

We base whether an individual received fertilizer on our tracking exercise. Whether an individual received information about fertilizer we base on our second household survey. We explore a wide set of characteristics. First are a set off characteristics related to the relationship between the ambassador and the receiver: family, field neighbors, discuss agriculture, physical distance between dwellings, and a receiver

centrality rank.⁴² We also look at characteristics of the receiver, specifically: age, gender, migrant status, and income. We separate out our investigation. That is, we explore for those that received from first-stage ambassadors then separately explore the characteristics of those that received from second-stage ambassadors. Finally, we look whether the network position of the first-stage ambassador had an impact on these results.

Column 1 in Error! Reference source not found. presents the results for first-stage ambassadors for receiving fertilizer while Column 2 reports results for receiving information about fertilizer. We find largely similar patterns for both resource and information. First-stage ambassadors are more likely to give fertilizer and information to people with whom they discuss agriculture. Someone being in their agricultural discussion network increases the chance of receiving fertilizer by 7%. For this group of ambassadors we do not find significant effects for blood ties and field neighbors. Additionally, geographic proximity plays a key role as ambassadors tend to share fertilizer with their geographical housing neighbors. Neighbors that live 1 kilometer closer are 12% more likely to receive fertilizer.⁴³ First-stage ambassadors are more likely to give fertilizer to community leaders (an increase of 3%). Initial recipients of goods may feel an obligation to include the chief in any perceived rewards. This might also explain they do not share with blood family, as opposed to 2nd stage ambassadors (see below). It also seems that first stage ambassadors share less to ethnic minorities.

Columns 3 to 4 in Error! Reference source not found. show the same results but for the second-stage mbassadors. Although there are similarities between the results related to those ambassadors to whom we seeded information and fertilizer directly (first-stage ambassadors) and those selected by them (second-stage ambassadors), there are differences. First, second-stage ambassadors are significantly more likely to give to individuals with whom they one of our social networks. For each that increases the probability by about 3%. Furthermore, second-stage ambassadors are not more likely to give to community leaders. In contrast, migrants are more likely to receive information from second stage ambassadors.

⁴² Within each community, we rank all possible receivers by centrality, excluding the ambassadors. This allows each ambassador to give to people of the maximum and minimum rank. This rank is normalized by dividing by the maximum rank.

⁴³ These results echo results by Sperling & Loevinsohn (1993) for neighboring Rwanda, who find that sharing of bean seeds was restricted to friends, family and neighbors.

Table 6 Predicting Diffusion Behavior

	(1) Fertilizer Use	(2) Fertilizer Info	(1) Fertilizer Use	(2) Fertilizer Info	(1) Fertilizer Use	(2) Fertilizer Info
	1 st stage	1 st stage	2 nd stage	2 nd stage	1 st stage	1 st stage
Ambassador						
Blood family	0.002 (0.016)	0.095 (0.203)	0.035*** (0.010)	0.014** (0.007)	0.000 (0.018)	0.012 (0.016)
Field Neighbors	0.032 (0.026)	0.204 (0.268)	0.033** (0.015)	0.023* (0.012)	0.027 (0.032)	0.020 (0.028)
Discuss Agriculture	0.071*** (0.026)	0.440 (0.267)	0.038*** (0.010)	0.024*** (0.008)	0.068** (0.030)	0.054* (0.029)
Physical Distance	-0.121*** (0.020)	-2.429*** (0.684)	-0.066*** (0.011)	-0.033*** (0.010)	-0.157*** (0.030)	-0.106** (0.045)
Receiver Centrality Rank	-0.026 (0.018)	-0.141 (0.308)	0.015 (0.010)	0.007 (0.009)	0.018 (0.028)	0.027 (0.027)
Receiver age	0.000 (0.000)	0.001 (0.004)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Receiver is female	-0.012 (0.010)	-0.045 (0.167)	-0.004 (0.005)	-0.002 (0.004)	-0.019 (0.016)	-0.013 (0.013)
Receiver Income index	0.002 (0.004)	0.044 (0.068)	0.003 (0.002)	-0.000 (0.002)	0.002 (0.005)	0.004 (0.005)
Same agricultural group	0.023 (0.020)	0.149 (0.266)	0.015 (0.009)	-0.003 (0.007)	0.009 (0.029)	0.005 (0.028)
Receiver is leader	0.027** (0.012)	0.240 (0.180)	-0.009 (0.007)	-0.006 (0.005)	0.039** (0.018)	0.010 (0.017)
Receiver is ethnic minority	-0.051* (0.031)	-0.148 (0.634)	-0.021 (0.014)	-0.001 (0.006)	-0.022 (0.036)	0.004 (0.050)
Receiver is migrant	-0.002 (0.008)	-0.067 (0.142)	0.008 (0.006)	0.009** (0.004)	-0.005 (0.014)	0.002 (0.013)
Treatment (1 = isolate ambassadors)					-0.043 (0.045)	-0.024 (0.053)
x Receiver Blood family					0.046 (0.038)	0.044 (0.038)
x Receiver Field Neighbor					0.016 (0.050)	0.064 (0.052)
x Receiver Discuss Agriculture					0.023 (0.053)	-0.034 (0.054)
x Receiver Physical Distance					0.056 (0.041)	-0.002 (0.053)
x Receiver Centrality Rank					-0.074** (0.036)	-0.064* (0.034)
x Receiver age					0.000 (0.000)	0.000 (0.000)
x Receiver is female					0.012 (0.021)	0.022 (0.018)
x Receiver is migrant					0.000 (0.007)	-0.003 (0.008)
x Receiver Income index					0.042 (0.036)	0.014 (0.031)
x Same agricultural group					-0.022 (0.024)	0.007 (0.023)
x Receiver is leader					-0.039 (0.056)	-0.008 (0.059)
x Receiver is ethnic minority					0.005 (0.017)	-0.012 (0.016)
Constant	0.150*** (0.021)	-2.104*** (0.375)	0.021** (0.009)	0.003 (0.007)	0.152*** (0.027)	0.104*** (0.028)
Observations	4644	4523	5155	5155	4644	4308

* p < 0.10, ** p < 0.05, *** p < 0.01.

Note: Analysis looks at all possible receivers for fertilizer and information givers. Dependent variable is 1 if respondents said they received fertilizer or information from first (second) stage ambassador. OLS Regressions. Robust standard errors in parentheses clustered at the ambassador and receiver level. Community fixed effects included
Income index is a normalized index (for the control group) calculated using the user-written Stata program WMEANEEFFECTS and comprises: # chickens owned, # goats or sheep owned, # cows owned, land size, land access.

Finally, in columns 5 to 6 of Error! Reference source not found. we explore whether these results are different by the type of first-stage ambassador. The only difference we find is that isolates are more likely to give to individuals with a lower centrality ranking. The person with the highest centrality ranking is 7% more likely to receive fertilizer than the person with the lowest centrality ranking. This implies that isolate ambassadors are giving to centrals, but central ambassadors do not give to isolates. This only holds with 90% significance for information. This is important if the goal is to have an equitable spread: giving to centrals is likely to concentrate the technology among the most central individuals. Furthermore, we see that isolates are more likely to share with individuals with whom no captured network connection exists. This final result is surprising, as it indicates that either isolates are sharing through unexpected interpersonal relationships that are not captured within our network questions or they are using the resource diffusion intervention as an opportunity to establish new connections. Our data is unable to untangle these two hypothesis.

In sum, we find several variables that can explain the distribution behavior of technology and information. Most importantly is geographic proximity, though the agricultural discussion network is also important. We find important distributional effects in that goods might cluster at the top centralities and not trickle down. There are clearly many other factors that explain distribution behavior, as the results are of moderate magnitude.

8 Conclusion

The diffusion of technologies and knowledge is a key component of political and economic development. While it is acknowledged that the identity of the initial recipients may impact resulting distribution patterns and development outcomes, few studies explicitly measure the consequences of their network position. We investigate whether the network position of initial recipients affects the use, knowledge willingness to pay for a technology, and the speed and spread of diffusion. We also explore whether attenuation of the

technology and associated information takes place. Since social networks may constrain who passes new information along to whom, we also consider whether the network position of the farmers affects *which* other community members receive the technology.

We implement a field experiment in forty communities in Eastern Congo. In an initial visit, we collect information about full social networks. In the subsequent visit we implemented an intervention in which selected ambassadors are trained on the importance and use of fertilizer. Subsequently, these ambassadors received packages of fertilizer and were asked to spread this fertilizer and their knowledge about fertilizer. We randomize whether these ambassadors are central or isolate, allowing us to estimate the effect of network position of initial recipients. In total, 225 transfers of fertilizer packs took place. During a third visit we conducted a tracking exercise and a second household survey to collect outcome measures.

We find that centrality measures based on family, farming, and agricultural discussion relationships are strongly correlated with observable characteristics. central farmers are more likely to hold a high status position in the community and interact with community leaders, are less likely to be a migrant, are in general richer, more literate, more likely to be male, and have larger families. We find no evidence of a treatment effect of centrality on the quantity of fertilizer use, knowledge or valuation. We also find no evidence that the spread is further. We do find the speed. We also find distributional effects. Network centrality affects *which* community members gain access to new technologies and learn about them. Both central and isolate farmers prefer sharing along existing social ties but the sharing behavior of central first-stage ambassadors versus isolate first-stage ambassadors differs in that isolate ambassadors tend to pass the technology on to more central individuals but the most central do not pass the technology to the most isolated. Our results have a number of implications.

First, our results indicate that resources and information pool to the center of networks regardless of entry-point when selection is based on centrality. Furthermore our results suggest that centrality is predictive of socioeconomic indicators of social prominence, with marginal households being more likely to sit on the periphery of their social networks. These two results taken in combination have important implications for intervention designs. In instances where practitioners have strong priorities on *who* is benefiting from their intervention and their focus is on marginal households, than designs relying on diffusion via central

individuals is inappropriate. Practitioners are better to directly distribute resources to these households as opposed to relying on social learning and diffusion dynamics.

Second, the strong attenuation of resources within the diffusion process highlights two important design considerations. The first is that if practitioners aim to leverage diffusion dynamics (as opposed to direct targeting), then considerations of periodic reinforcement of knowledge and resources may strengthen ultimate diffusion by re-legitimizing and re-invigorating the diffusion process. Our study shows that quality of information and rate of adoption decreases with each step away from the initial injection point. Designing programs that incorporate multiple practitioner-community interactions may help alleviate some of these inherent attenuation effects.

Finally, our results speak to the cost-benefit trade-off of network entry-point selection and resulting diffusion outcomes of reach, speed, and benefit distribution. Development project designs must take into account network sizes, desired spread and speed of diffusion, and intended recipients when selecting entry points. With the goal of reducing implementation costs of introducing new technologies and information while maximizing overall reach, the use of social network analysis is called into question. However, measures of network centrality could be useful for attaining goals of reaching household on the periphery of social networks, although greater research into whether more observable correlates of ethnicity or migrant status could be used as inexpensive alternatives of identifying marginalized households. In either case, as the collection of network data is costly, the ability to predict diffusion by social network analysis is beneficial contingent on leveraging network-gains through less-expensive alternative measures. Our study contributes to existing literature exploring this trade-off by quantifying potential gains of network-targeting and exploring relationships between network metrics and more observable socioeconomic indicators in identifying cost-feasible designs of network-targeting for development practitioners. Research on more cost-effective designs based on these network-socioeconomic relationships or other selection methods (e.g. geo-spatial networks; community nomination processes) is gaining momentum within a developing country context (Banerjee et al., 2013b; Beaman et al., 2015; Kim et al., 2015; Nourani, 2016). However significant gaps remain and further theoretical modeling and empirical testing is required before concrete policy recommendations are definitive.

References

- Agyei-Holmes, A., Ayerakwa, H. M., Osei, R. D., & Osei-Akoto, I. (2011). *Training and Farmer Productivity : An Evaluation using RCT for the MCA-Ghana Programme*.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2010). Preference heterogeneity in experiments: Comparing the field and laboratory. *Journal of Economic Behavior and Organization*, 73(2), 209–224. <https://doi.org/10.1016/j.jebo.2009.09.006>
- Anderson, J. R., & Feder, G. (2007). Chapter 44 Agricultural Extension. *Handbook of Agricultural Economics*, 3(6), 2343–2378. [https://doi.org/10.1016/S1574-0072\(06\)03044-1](https://doi.org/10.1016/S1574-0072(06)03044-1)
- Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America*, 106(51), 21544–21549. <https://doi.org/10.1073/pnas.0908800106>
- Aral, S., Muchnik, L. E. V., & Sundararajan, A. (2013). Engineering social contagions: Optimal network seeding in the presence of homophily. *Network Science*, 1(2), 125–153. <https://doi.org/10.1017/nws.2013.6>
- Aral, S., & Walker, D. (2012). Identifying Influential and Susceptible Members of Social Networks. *Science*, 337, 337–341. <https://doi.org/10.1126/science.1215842>
- Bailard, C. S. (2010). Mobile Phone Diffusion and Corruption in Africa. *Liberation Technologies Seminar Series*. <https://doi.org/10.1080/10584600903053684>
- Bandiera, O., & Rasul, I. (2006). SOCIAL NETWORKS AND TECHNOLOGY ADOPTION. *The Economic*, 116(514), 869–902.
- Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013a). The diffusion of microfinance. *Science*, 341, 363–370. <https://doi.org/10.1126/science.1236498>
- Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013b). The diffusion of microfinance. *Science*, 341(6144), 1236498. <https://doi.org/10.1126/science.1236498>
- Banerjee, A. V., & Duflo, E. (2009). The Experimental Approach to Development Economics. *Annual Reviews of Economics*, 1, 151–178. <https://doi.org/10.1146/annurev.economics.050708.143235>
- Barrett, C. B., & Carter, M. R. (2010). The power and pitfalls of experiments in development economics: Some non-random reflections. *Applied Economic Perspectives and Policy*, 32(4), 515–548. <https://doi.org/10.1093/aep/pqp023>
- Beaman, L., Benyishay, A., William, C., & Mobarak, A. M. (2015). *Can Network Theory based Targeting Increase Technology Adoption ?*
- Benjamin, D. (1992). Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models. *Econometrica*, 60(2), 287–322.
- BenYishay, A., & Mobarak, A. M. (2014). Social Learning and Communication. *NBER Working Paper*, (20139). <https://doi.org/10.3386/w20139>
- Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist*, 37(3).
- Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, 2(1), 113–120.
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29, 555–564. <https://doi.org/10.1016/j.socnet.2007.04.002>
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27, 55–71. <https://doi.org/10.1016/j.socnet.2004.11.008>
- Borgatti, S. P., Carley, K. M., & Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, 28, 124–136. <https://doi.org/10.1016/j.socnet.2005.05.001>
- Bulte, E., Leuvel, K., Nillesen, E. E. M., Pieters, J., & Voors, M. (2015). *Farm Households in Eastern Congo, Baseline Survey Report*.
- Bursztyn, L., Ederer, F., Ferman, B., & Yuchtman, N. (2014). Understanding Mechanisms Underlying Peer Effects: Evidence from a Field Experiment on Financial Decisions. *Econometrica*, 82(4), 1273–1301.
- Cai, J. (2012). Social Networks and the Decision to Insure : Evidence from Randomized Experiments in China *.
- Cai, J., Janvry, A. De, & Sadoulet, E. (2015). Social Networks and the Decision to Insure (Supplementary Figures and Tables). *American Economic Journal: Applied Economics*, 7(2), 81–108. <https://doi.org/10.2139/ssrn.2161686>
- Carter, M. R., Laajaj, R., & Yang, D. (2014). Subsidies and the persistence of technology adoption: Field experimental evidence from Mozambique. *Nber, w20465*. <https://doi.org/10.3386/w20465>
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329, 1194–1197. <https://doi.org/10.1126/science.1185231>
- Centola, D. M., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3), 702–734. <https://doi.org/10.1086/521848>
- Cerdán-Infantes, P. (2008). The impact of agricultural extension services: The case of grape production in Argentina. *OVE Working Papers*. Retrieved from <http://www.iadb.org/intal/intalcdi/PE/2009/03159.pdf>
- CGIAR. (2014). *Project Cialca DR Congo, Rwanda, Burundi*.

- Chami, G. F., Ahnert, S. E., Voors, M. J., & Kontoleon, A. A. (2014). Social network analysis predicts health behaviours and self-reported health in African villages. *PLoS ONE*, 9(7). <https://doi.org/10.1371/journal.pone.0103500>
- Chandrasekhar, A. G., & Lewis, R. (2010). *Econometrics of sampled networks*. *Networks*. <https://doi.org/10.1017/CBO9781107415324.004>
- Choi, H., Kim, S.-H., & Lee, J. (2010). Role of network structure and network effects in diffusion of innovations. *Industrial Marketing Management*, 39(1), 170–177. <https://doi.org/10.1016/j.indmarman.2008.08.006>
- Christiaensen, L., Pan, L., & Wang, S. (2010). *Drivers of Poverty Reduction in Lagging Regions: Evidence from rural Western China* (No. 2010,35). <https://doi.org/10.1007/s10273-011-1262-2>
- Chuang, Y., & Schechter, L. (2000). Social Networks in Developing Countries. *Annual Review of Resource Economics*, (7). <https://doi.org/10.1146/annurev-resource-100814-125123>
- CIALCA. (2007). *Progress Report December 2007. Presentations*.
- Cilliers, J., Kasirye, I., Leaver, C., Serneels, P., & Zeitlin, A. (2014). *Improving Teacher Attendance Using a Locally Managed Monitoring Scheme : Evidence From Ugandan Primary Schools* (Policy Note No. 14/0189).
- Coghlan, B., Ngoy, P., Mulumba, F., Hardy, C., Bemo, V. N., Stewart, T., ... Brennan, R. (2007). *Mortality in the Democratic Republic of Congo: An Ongoing Crisis*. *International Rescue Committee*. Retrieved from http://www.rescue.org/sites/default/files/migrated/resources/2007/2006-7_congomortalitysurvey.pdf
- Conley, B. T. G., & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *The American Economic Review*, 100(1), 35–69.
- Crawford, E. W., & Kelly, V. a. (2006). *Alternative Approaches for Promoting Fertilizer Use in Africa* (Agriculture and Rural Development Discussion Paper No. 22).
- Cunguara, B., & Moder, K. (2011). Is Agricultural Extension Helping the Poor? Evidence from Rural Mozambique. *Journal of African Economics*, 20(4), 562–595.
- Dercon, S., & Christiaensen, L. (2008). *Consumption Risk, Technology Adoption, and Poverty Traps: Evidence from Ethiopia*. (World Economy & Finance Working Paper Series No. WEF 0035).
- Dixit, A., & Pindyck, R. (1994). *Investment Under Uncertainty*. *Investment Under Uncertainty*. Princeton University Press. <https://doi.org/10.2307/1909571>
- Druilhe, Z., & Barreiro-hurlé, J. (2012). *Fertilizer subsidies in sub-Saharan Africa* (ESA Working Paper No. 12–4). Retrieved from www.fao.org/economic/esa
- Duflo, E. (2006). Field Experiments in Development Economics 1. *Econometric Society Monographs*, 42(12), 1–30. Retrieved from <http://qed.econ.queensu.ca/pub/students/phds/khana/econ239/readings/duflo.pdf>
- Duflo, E., Kremer, M., & Robinson, J. (2011). How high are rates of return to fertilizer? Evidence from field experiments in Kenya. *The American Economic Review*, 101(6), 2350–2390. <https://doi.org/10.1257/aer.98.2.482>
- Dutta, S., Lanvin, B., & Paua, F. (2004). *The Global Information Technology Report 2003-2004 Towards an Equitable Information Society*.
- Ellis, F. (1993). *Peasant Economics: farm household in agrarian development* (2nd ed.). Cambridge University Press.
- Emerick, K., Janvry, A. De, Sadoulet, E., & Dar, M. H. (2016). *Identifying early adopters , enhancing learning , and the diffusion of agricultural technology*.
- FAO. (2000). *Fertilizers and Their Use*. *Tennessee University Extension*. FAO Rome. [https://doi.org/10.1016/S0022-3913\(12\)00047-9](https://doi.org/10.1016/S0022-3913(12)00047-9)
- FAO. (2009). *The state of food and agriculture: Livestock in balance*. *Food and Agriculture Organisation of the United Nations*. [https://doi.org/10.1016/S0140-6736\(75\)92740-3](https://doi.org/10.1016/S0140-6736(75)92740-3)
- FAO. (2015). *FAO Statistical Pocketbook 2015*. *Food and Agriculture Organization of the United Nations*. <https://doi.org/978-92-5-108802-9>
- Feder, G., Anderson, J. R., Birner, R., & Deininger, K. (2010). Promises and Realities of Community-Based Agricultural Extension. *Community, Market and State in Development*, (March), 187–210. https://doi.org/10.1057/9780230295018_12
- Feder, G., Murgai, R., & Quizon, J. B. (2003). *Sending Farmers Back to School: The Impact of Farmer Field Schools in Indonesia* (No. 3022). *Policy Research Working Paper*. Retrieved from <http://econ.worldbank.org>
- Feder, G., & Slade, R. (1984). Information Acquisition and the Adoption of New Technology. *American Journal of Agricultural Economics*, 66(3), 312–320. <https://doi.org/10.1287/mnsc.31.11.1372>
- Feder, G., & Umali, D. L. (1993). The Adoption of Agricultural Innovations. *Technological Forecasting and Social Change*, 43, 215–239.
- Ferment, a. M., van Asten, P. J. a, Tittone, P., van Wijk, M. T., & Giller, K. E. (2009). Closing the cassava yield gap: An analysis from smallholder farms in East Africa. *Field Crops Research*, 112(1), 24–36. <https://doi.org/10.1016/j.fcr.2009.01.009>
- Foltz, J. D., Aldana, U. T., & Laris, P. (2012). *The Sabel's Silent Maize Revolution: Analyzing Maize Productivity in Mali at the Farm Level*. *National Bureau of Economic Research Working Paper Series* (Vol. No. 17801). <https://doi.org/10.3386/w17801>
- Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of Technology Adoption. *Annual Review of Economics*, 2. <https://doi.org/10.1146/annurev.economics.102308.124433>
- Gneezy, U., & Imas, A. (2017). Lab in the Field: Measuring Preferences in the Wild. *Handbook of Field Experiments*, *Forthcoming*, 1–40. <https://doi.org/10.1017/CBO9781107415324.004>

- Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., ... Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science*, 327, 812–818. <https://doi.org/10.1126/science.1185383>
- Grossman, G., Humphreys, M., & Sacramone-Lutz, G. (2014). “I wld like u WMP to extend electricity 2 our village”: On Information Technology and Interest Articulation. *American Political Science Review*, 108(3), 688–705. <https://doi.org/10.1017/S0003055414000331>
- Haider, R., Ashworth, A., Kabir, I., & Huttly, S. R. (2000). Effect of Community-Based Peer Counsellors on Exclusive Breastfeeding Practices in Dhaka, Bangladesh: A Randomised Controlled Trial. *Lancet*, 356(9242), 1643–1647.
- Hanna, R., Schwartzstein, J., Benjamin, D., Gentzkow, M., Halac, M., Newman, A., & Ogden, T. (2014). Learning Through Noticing: Theory and Experimental Evidence in Farming.
- Hardy, M., & Mccasland, J. (2015). *It Takes Two : Experimental Evidence on the Determinants of Technology Diffusion*.
- Harrison, G. W., & List, J. a. (2004). Field Experiments. *Journal of Economic Literature*, 42(4), 1009–1055.
- Hiebert, D. L. (1974). Risk, Learning, and the Adoption of Fertilizer Responsive Seed Varieties. *American Journal of Agricultural Economics*, 56(4), 764–768.
- Irz, X., Lin, L., Thirtle, C., & Wiggins, S. (2001). Agricultural Productivity Growth and Poverty Alleviation. *Development Policy Review*, 19(4), 449–466. <https://doi.org/10.1111/1467-7679.00144>
- Janvry, A. De, Fafchamps, M., & Sadoulet, E. (1991). Peasant Household Behaviour with Missing Markets : Some Paradoxes Explained. *The Economic Journal*, 101(409), 1400–1417.
- Just, R. E., & Zilberman, D. (1983). Stochastic Structure , Farm Size and Technology Adoption in Developing. *Oxford Economic Papers*, 35(2), 307–328.
- Kenzior, J., Zibika, J. P., & Voors, M. (2015). *Social relationships, local institutions, and the diffusion of improved variety seed and field management techniques in rural communities: six case studies in South Kivu, DRC*.
- Key, N., Sadoulet, E., & Janvry, A. De. (2000). Transactions Costs and Agricultural Household Supply Response. *American Journal of Agricultural Economics*, 82(2), 245–259. <https://doi.org/10.1111/0002-9092.00022>
- Kim, D. A., Hwong, A. R., Stafford, D., Hughes, D. A., O'Malley, A. J., Fowler, J. H., & Christakis, N. A. (2015). Social network targeting to maximise population behaviour change: A cluster randomised controlled trial. *The Lancet*, 386, 145–153. [https://doi.org/10.1016/S0140-6736\(15\)60095-2](https://doi.org/10.1016/S0140-6736(15)60095-2)
- Kimbrough, E., Smith, V. L., & Wilson, B. J. (2008). Historical Property Rights, Sociality, and the Emergence of Impersonal Exchange in Long-distance Trade. *American Economic Review*, 98(3).
- Kincaid, D. L. (2000). Social Networks, ideation, and contraceptive behavior in Bangladesh: A longitudinal analysis. *Social Science & Medicine*, 50(2).
- Kondylis, F., Mueller, V., & Zhu, S. J. (2014). *Seeing is believing? Evidence from an extension network experiment* (Policy Research Working Paper No. 7000).
- Lambrecht, I., Vanlauwe, B., & Maertens, M. (2016). Agricultural extension in Eastern Democratic Republic of Congo: does gender matter? *European Review of Agricultural Economics*, jbv039. <https://doi.org/10.1093/erae/jbv039>
- Levitt, S. D., & List, J. A. (2007). About the Real World ? *Journal of Economic Perspectives*, 21(2), 153–174.
- Maass, B. L., Musale, D. K., Chiuri, W. L., Gassner, A., & Peters, M. (2012). Challenges and opportunities for smallholder livestock production in post-conflict South Kivu, eastern DR Congo. *Tropical Animal Health and Production*, 44(6), 1221–1232. <https://doi.org/10.1007/s11250-011-0061-5>
- Maertens, A., & Barrett, C. B. (2012). Measuring social networks' effects on agricultural technology adoption. *American Journal of Agricultural Economics*, (649330).
- Marenja, P. P., & Barrett, C. B. (2009). State-conditional fertilizer yield response on Western Kenyan Farms. *American Journal of Agricultural Economics*, 91(4), 991–1006. <https://doi.org/10.1111/j.1467-8276.2009.01313.x>
- Morris, M., Kelly, V. a, Kopicki, R. J., & Byerlee, D. (2007). *Fertilizer uses in African Agriculture: Lessons Learned and Good Practice Guidelines*. Washington D.C.: The World Bank.
- Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *The Quarterly Journal of Economics*, (May), 549–599. <https://doi.org/10.1162/003355303321675455>
- Nisrane, F., Berhane, G., Asrat, S., Getachew, G., Taffesse, A. S., & Hoddinott, J. (2011). *Sources of Inefficiency and Growth in Agricultural Output in Subsistence Agriculture : A Stochastic Frontier Analysis* (No. 19). *Ethiopia Strategy Support Program (ESSP II)*. Retrieved from WWW.ifpri.org
- Nourani, V. (2016). *Social Network Effects of Technology Adoption : Investing with Family , Learning from Friends & Reacting to Acquaintances*.
- Olken, B. A. (2006). Corruption and the costs of redistribution: Micro evidence from Indonesia. *Journal of Public Economics*, 90(4–5), 853–870. <https://doi.org/10.1016/j.jpubeco.2005.05.004>
- Oster, E., & Thornton, R. (2012). Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-Up. *Journal of the European Economic Association*, 10(6), 1263–1293. <https://doi.org/10.1111/j.1542-4774.2012.01090.x>
- Pentland, A. (2014). *Social Physics: How good ideas spread-the lessons from a new science*. Penguin.
- Pypers, P., Sanginga, J. M., Kasereka, B., Walangululu, M., & Vanlauwe, B. (2011). Increased productivity through integrated soil fertility management in cassava-legume intercropping systems in the highlands of Sud-Kivu, DR Congo. *Field Crops Research*, 120(1), 76–85. <https://doi.org/10.1016/j.fcr.2010.09.004>
- Reinikka, R., & Svensson, J. (2005). FIGHTING CORRUPTION TO IMPROVE SCHOOLING : EVIDENCE

- FROM A NEWSPAPER CAMPAIGN IN UGANDA. *Journal of the European Economic Association*, 3(2), 259–267.
- Roe, B., & Just, D. (2009). Internal and External Validity in Economics Research: Tradeoffs between Experiments, Field Experiments, Natural Experiments, and Field Data. *American Journal of Agricultural Economics*, 91(5), 1266–1271.
- Rogers, E. M. (1995). *Diffusion of Innovations*.
- Rosegrant, M. W., & Cline, S. a. (2003). Global Food Security: Challenges and Policies. *Science*, 302, 1917–1919. <https://doi.org/10.1126/science.1092958>
- Simpson, B. M., Franzel, S., Degrande, A., Kundhlande, G., & Tsafack, S. (2015). Farmer-to-Farmer Extension: Issues in Planning and Implementation. *MEAS Technical Note*, (May).
- Smith, A. (1759). *The Theory of Moral Sentiments*.
- Smith, R. D. (2006). It's not just what you do, it's the way that you do it: The effect of different payment card formats and survey administration on willingness to pay for health gain. *Health Economics*, 15(3), 281–293. <https://doi.org/10.1002/hec.1055>
- Taye, H. (2013). Evaluating the impact of agricultural extension programmes in sub-Saharan Africa: Challenges and prospects. *African Evaluation Journal*, 1(1), 1–9. <https://doi.org/10.4102/aej.v1i1.19>
- Taylor, J. E., & Adelman, I. (2003). Agricultural Household Models : Genesis , Evolution , and Extensions. *Review of Economics of the Household*, 1(1).
- Tittonell, P., & Giller, K. E. (2013). When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research*, 143, 76–90. <https://doi.org/10.1016/j.fcr.2012.10.007>
- UNDP. (2013). *2013 Human Development Report: The Rise of the South: Human Progress in a Diverse World*.
- Valente, T. W. (1996). Network models of the diffusion of innovations. *Computational & Mathematical Organization Theory*, 2(2), 163–164.
- Valente, T. W. (2012). Network Interventions. *Science*, 337(6090), 49–53. <https://doi.org/10.1126/science.1217330>
- Valente, T. W., & Davis, R. L. (1999). Accelerating the Diffusion of Innovations using Opinion Leaders. *The ANNALS of the American Academy of Political and Social Science*. <https://doi.org/10.1177/0002716299566001005>
- Valente, T. W., & Pumpuang, P. (2007). Identifying Opinion Leaders to Promote Behavior Change. *Health Education & Behavior*, 34(6), 881–896. <https://doi.org/10.1177/1090198106297855>
- Vlassenroot, K., & Huggins, C. (2005). Land, Migration and Conflict in eastern DRC. In *From the ground up: land rights, conflict and peace in sub-Saharan Africa* (pp. 115–195).
- Walsh, G., Gwinner, K. P., & Swanson, S. R. (2004). What makes mavens tick? Exploring the motives of market mavens' initiation of information diffusion. *Journal of Consumer Marketing*, 21(2), 109–122. <https://doi.org/10.1108/07363760410525678>
- Woomer, P., Huising, J., Giller, K. E., & Et, A. (2014). *N2Africa : Final Report of the First Phase 2009-2013*. Retrieved from www.N2Africa.org
- WORLD, B. (2008). *Agriculture for Development*.
- WORLD, B. (2015). World Bank Development Indicators. Retrieved from data.worldbank.org/indicator/NY.GDP.PCAP.CD
- Zingore, S. (2011). Maize productivity and response to fertilizer use as affected by soil fertility variability manure application and cropping system in Africa. *Better Crops*, 95(1), 9.

Appendix

Table 7. Attrition

	(1) Panel Attrition	(2) Absence
Age	-0.005 (0.005)	0.007 (0.010)
Female	0.303 (0.236)	0.272 (0.218)
Centrality	0.075 (0.283)	
War Participation	0.302 (0.424)	
War Exposure	-0.183** (0.083)	
Household Size	0.005 (0.022)	
Migrant	-0.015 (0.200)	-1.233*** (0.409)
Fertilizer Knowledge (baseline)	-0.060 (0.046)	
Land Size	-0.001 (0.001)	
Student		0.026 (0.998)
Farmer		-1.244*** (0.238)
Wage Labourer		-0.530 (0.420)
Petty Trader		0.794** (0.334)
Miner		-0.030 (0.310)
Unemployed		-1.299* (0.787)
Other Work		-0.356 (0.470)
Constant	-1.763*** (0.478)	-2.519*** (0.547)
Observations	2461	2330
Attrition	0.091	
Absence		0.031

p < 0.10, ** p < 0.05, *** p < 0.01.

Robust standard errors in parentheses clustered at the community level.

CHAPTER 4

Community Selection of Targets in Network Diffusion Interventions

Abstract

Interventions aiming to increase adoption rates through information campaigns often rely on network diffusion strategies. A critical design component within these initiatives is selection of the network entry points in order to facilitate faster and/or wider information spread. This study combines a community-wide polling of network entry-points combined with detailed community network and socio-economic data. First we explore what attributes are prioritized by community members in nominating a resident farmer as an extension contact-point. Second, we use simulations to compare the diffusion spread of top-nominated individuals as network entry-points compared to entry-points that achieve maximal spread within diffusion simulations. We find that community members prioritize network connectedness, pro-social preferences, and socioeconomic indicators of gender, age, formal leadership, and education levels within their nomination decisions. Furthermore, receiving the top three most amount of nominations is found to be significantly correlated with selection as an optimal entry-point within the diffusion simulation.

Publication Status: Ross, M., P. Hofman, P. van der Windt, M. Voors. 2017. “Selection of Targets in Network Diffusion Interventions.” Working Paper.

1 Introduction

Adoption of growth enhancing technologies and behaviors is essential for development (Alvarez et al., 2013). Technology adoption has been studied in a wide range of applications including agricultural production (Beaman et al, 2015; Emerick et al., 2016), health (Oster and Thornton, 2012), finance (Banerjee et al., 2013; Breza and Chandrasekhar, 2015; Chuang and Schechter, 2015), labor (Beaman, Keleher and Magruder, 2013) and institutions (Breza et al., 2014). Information on the returns to investment for new innovations is a key component of adoption decisions. Interventions aiming to increase adoption rates through information campaigns often rely on network diffusion strategies. A critical design component within these initiatives is selection of the network entry points in order to facilitate faster and/or wider information spread. Social network diffusion dynamics have gained attention in the previous decades as offering an effective means for identifying optimal network contact-points based on diffusion objectives (Aral & Walker, 2012; D. Centola, 2010). However network-based approaches are data-intensive and require detailed information on underlying network structures. For development initiatives, the data demands of such approaches can be prohibitively high.

For technology adoption, network-based studies focus largely on the underlying network structure and the position of contact-points within the structure. However the role of an individual's social preferences, incentives, motivations, and prior beliefs on the innovation has also been highlighted as important for an individual's social influence within technology spread (Pentland, 2014; T. W. Valente & Pumpuang, 2007; Van Eck, Jager, & Leeflang, 2011). Therefore optimal selection of network injection-points must account for both network factors and personal traits when selecting contact-points, although the two factors are highly interdependent.

In the context of agricultural innovation adoption, farmer-led extension programs have become a popular tool for raising the adoption rates of productivity-enhancing inputs. Evidence suggests that technology adoption rates could be increased through network based targeting within agriculture (Beaman et al., 2015). With the high costs of network-data, network-based targeting is often unrealistic as a widespread policy tool in many contexts. Community members have incomplete snapshots of both their social network structures (Banerjee et al., 2014) and their peer's relevant personal traits. Aggregating each community member's information set provides a detailed, but incomplete, overview of the key selection considerations.

Full community-nomination of contact farmers may therefore offer a less data-intensive alternative selection method.

This study contributes to the literature in two ways. First we explore what attributes are prioritized by community members in nominating a resident farmer as an extension contact-point. This analysis explores whether community members are incorporating network and personal characteristics relevant to improved diffusion processes into their nomination decisions. We find that community members prioritize network connectedness, pro-social preferences, and socioeconomic indicators of gender, age, formal leadership, and education levels within their nomination decisions. Second, we use simulations to compare the diffusion spread of top-nominated individuals as network entry-points compared to entry-points that achieve maximal spread within diffusion simulations. Receiving the top three most amount of nominations is found to be significantly correlated with selection as an optimal entry-point within the diffusion simulation. Given the cost-constraints of development practitioners, our results suggest that the trade-off between data-demands and gains in diffusion makes community-wide nomination a practical and effective alternative to network-based selection of contact points.

The remainder of the paper is organized as follows. Section 2 covers the existing literature on network diffusion, technology adoption, and agricultural extension. Section 3 introduces the project design and local context of rural communities in Eastern DRC. Section 4 outlines our empirical strategy and provides summary statistics of the data used in analysis. Section 6 catalogues our empirical results and section 7 concludes.

2 Literature

2.1 Targeting for Diffusion

An individual is said to be socially influential if their behaviors or beliefs affect the beliefs, and ultimately adoption decisions, of peers. Thus the more influential an individual is, the more weight people give to information received from him or her (Leonard-Barton, 1985; T. W. Valente & Pumpuang, 2007). Social influencers are thought to catalyze the diffusion of innovations and reduce lag time between introduction of innovations and critical mass thresholds for cascades being achieved (Davis, 1999; Thomas W. Valente, 1996). For interventions attempting to achieve high adoption rates, network injection points should be

individuals that are both socially influential and well positioned within the network structure for efficient information spread.

In identifying socially influential individuals, no clear pattern of socioeconomic traits has been found that distinguishes network influencers from non-influencers (Aral & Walker, 2012; Chan & Misra, 2013). However social influence within technology adoption has been found to increase with topic expertise, comparability in conditions relevant to the innovation, and pro-social values (Chan & Misra, 2013; Walsh, Gwinner, & Swanson, 2004). Additionally, socially influential individuals are found to have more central positions within their social network structures (Van Eck et al., 2011). An individual's social influence will therefore vary both by domain and network links. Social influence captures the credibility given to information received by an individual, but network positioning dictates the extent to which information will diffuse onward through the network structure. Considerable attention has been focused on determining how diffusion patterns vary by selecting network contact points using different measures of connectedness (Cho, Hwang, & Lee, 2012).

Different measures of connectedness capture different types of information flows (Borgatti, 2005). Eigenvector centrality is one measure of social connectedness that is considered a strong approximation of an individual's probability of participating in network information flows. This is due to it imposing no restrictions on neither the frequency with which information passes nor the number of pathways by which it travels (Borgatti, 2005; Bonacich 1987, 1991). For this reason, we prioritize eigenvector centrality as the network positioning parameter that captures an individual's level of expected influence within their given network. In an online social network experiment, Centola (2010) found that behavior spreads faster across networks with a clustered-lattice structure compared to random networks. We disaggregate this finding to the individual level with a node-specific measure of clustering, which measures the proportion of mutual connections between each pair of nodes within a network, as a second proxy for an individual's potential level of influence and information diffusion capabilities within their network.

In simulating technology adoption processes at the network level, we focus on threshold models. This choice is based on these models of diffusion offering a parsimonious method for theoretically estimating entry points that maximize spread. Maximal spread is defined as the highest proportion of network members adopting the new innovation. These models are built on the assumption that any given individual will adopt

if exposed to a threshold number of prior adopters (Acemoglu & Ozdaglar, 2011; D. M. Centola & Macy, 2007; Granovetter, 1978). Adjusting the minimum threshold requirement for adoption will result in diffusion patterns variations.

If exposure to a single adopter is sufficient for an individual to also adopt, then diffusion patterns will be similar regardless of entry point chosen assuming a long enough time horizon. These ‘simple contagion’ diffusions apply only to technologies that have very low levels of associated knowledge and easily realizable returns to adoption. For innovations that entail greater knowledge or greater uncertainty on returns to adoption, multiple exposures are required to spur adoption. This higher threshold case is termed ‘complex contagion’ (Centola & Macy, 2007). Technologies that more closely resemble complex contagion diffusion patterns, practitioners are best off selecting socially-close entry points with shared connections to ensure minimum exposure thresholds are reached in the early stages of diffusion (Beaman, Benyishay, William, & Mobarak, 2015). We use complex contagion simulation models in modeling the diffusion of agricultural innovations given the higher knowledge-intensity and uncertainty inherent in agricultural technologies.

2.2 Agricultural Extension and Contact-farmer Selection

Agricultural output within sub-Saharan Africa continues to fall short of estimated levels of potential production. Agricultural extension programs have gained immense popularity in recent decades as a policy tool for addressing persistently low smallholder productivity levels and issues of food security within the developing world. In recent years, more participatory farmer-led extension programs have rapidly gained popularity despite limited evidence on their effectiveness (Anderson & Feder, 2004; Feder et al., 2010; Feder, Murgai, & Quizon, 2003). In farmer-led agricultural extension programs, a few farmers from each participating community are selected to act as the contact point between extension agents and the community. These selected community members are then responsible for ferrying information between the extension agents and their communities (Kondylis, Mueller, & Zhu, 2014). These farmer-led designs are motivated by the belief that community farmers have in-depth knowledge of local agri-climatic and market conditions, production processes, and are known to their fellow community members. These advantages are thought to make contact-farmers better able to present information and technologies to their communities in ways that are relatable, confidence-inducing, and cost-effective compared to external extension professionals (Kiptot and Franzel, 2014).

Extension programs vary in their designs but typically include training sessions, demonstration plots, one-on-one assistance. Almost always the programs rely on word-of-mouth communication through interpersonal relationships to spread information on improved agricultural methods and inputs (Lukuyu et al, 2012; Franzel et al., 2014). Farmer-led extension designs are intended to improve the diffusion of agricultural knowledge and innovation by utilizing individuals with high social influence as the contact point between extension professionals and respective communities.

Contact-farmers are typically jointly selected by communities and extension practitioners with varying levels of involvement by each. Studies have explored both motivations of contact-farmers in carrying out their extension responsibilities and what traits communities found desirable in selecting contact-farmers. Typically contact-farmers were found to not receive any financial compensation for their extension roles. Instead they report being motivated by altruism, access to knowledge, opportunities to experiment with new resources, and potential for personal networking (Simpson, Franzel, Degrande, Kundhlande, & Tsafack, 2015). On the other side, communities reported that pro-social behaviors and agricultural progressiveness were key determinants in selecting peers as contact-farmers to represent them (Lukuyu et al., 2012). Finally, studies have explored relational differences between contact-farmers and their community members. Contact-farmers who were only moderately better performing than community averages were most effective in diffusing information compared to those who displayed large gains in behavior performance

In a study on a pest management extension project, those contact-farmers who were more productive than their information recipients, but only moderately, were most effective in the transmission of knowledge. Contact-farmers with excessive differences in performance and socio-economic indicators are thus found to be less effective in influencing community adoption rates (BenYishay & Mobarak, 2014; Feder & Savastano, 2006)

3 Study Design

3.1 Regional Context

Our study is set within the South Kivu province of eastern DRC. The region is characterized by underdeveloped infrastructure, low agricultural productivity, and armed conflict instability (Coghlan et al., 2007; Huggins, 2010; Vlassenroot & Raeymaekers, 2004; World Bank, 2015) Currently, the use of fertilizer

within the region is very low; only 3% of farmers use fertilizer and average quantities applied are 0.3 kilograms per hectare, compared to 14 Kg/Ha for sub-Saharan Africa as a whole and 166 Kg/Ha average application in Asia (Bulte et al., 2015; FAO, 2015). Poor transport infrastructure and limited access to input markets cause production-enhancing technologies to be under-utilized (Pypers et al., 2011).

Within the study-region, economic and social interactions are predominantly local due to the underdeveloped transport infrastructure and hilly landscape. Most of the agricultural activities and farmer groups are formed within the community. As a result, a person's social network position is pivotal in accessing and controlling information and resources. Relationships of family and geographical field and house neighbors were found to be of particularly importance for agricultural resource access and spread (Kendzior, Zibika, & Voors, 2015; Sperling & Loevinsohn, 1993).

3.2 Data

Our study was conducted in rural communities in South Kivu province in eastern Democratic Republic of Congo. We selected 40 small communities with a maximum size of 100 households. In each community, we first conducted a full household census and then interviewed each household head⁴⁴, recording details on demographics, agricultural production information from the most recent production season, and social preferences. In addition, we asked for each social network tie to other community members (see section 3.3 for details). In interviewing the chief, we also asked about access to markets and the history of land conflicts in the past year. **Table 1** displays our descriptive data and appendix Table A1 includes variable definitions.

⁴⁴ When household heads were not available (due to seasonal migration or prolonged travel from the community), a replacement within the household was surveyed. Replacements were asked to answer from the perspective of the household head for all survey questions.

Table 1 Summary Statistics at Community Level and Household Level of Survey Indicators

VARIABLES	(1) N	(2) Mean	(3) S.D.	(4) Min	(5) Max
Community Level					
Community Indicators					
Number of Households	40	64.08	16.08	23	81
Number of Land Conflicts	40	1.25	2.468	0	12
Market Transport Cost (Francs)	40	1,780	1,135	0	4,000
Years as Chief	40	26.48	15.83	1	65
Community Networks					
Centrality	40	0.341	0.0826	0.179	0.547
Clustering Coefficient	40	0.612	0.186	0.334	1.028
Household Level					
Network Indicators					
Eigenvector Centrality	2,533	0.328	0.263	0	1
Clustering Coefficient	2,537	0.384	0.242	0	1
Social Preferences					
Other	2,475	0.259	0.438	0	1
Selfish	2,475	0.0606	0.239	0	1
Inequality Averse	2,475	0.592	0.492	0	1
Generous	2,475	0.0877	0.283	0	1
Amount Shared	2,475	1.265	0.422	0	3
Household Indicators					
Native Language	2,563	0.854	0.353	0	1
Female (=1)	2,553	0.322	0.467	0	1
Age (yrs.)	2,553	45.9	17.53	12	103
Household Size	2,475	6.334	3.161	1	30
Migrant (=1)	2,554	0.327	0.469	0	1
<i>Highest Education</i>					
Primary	2,475	0.311	0.463	0	1
Secondary	2,475	0.232	0.422	0	1
Tertiary	2,475	0.0182	0.134	0	1
War Exposure	2,475	2.312	1.04	0	4.5
Frequency Talk Chief	2,428	6.287	8.31	0	31
Community Leader	2,475	0.316	0.465	0	1
Wealth Index	2,390	-0.0189	0.953	-2.249	12.93
Agricultural Indicators					
Farm Experience (yrs.)	2,473	30.87	17.54	0	90
Primary Farmer	2,563	0.767	0.423	0	1
Fertilizer Knowledge	2,465	1.268	1.453	0	6.5
Labor Collective	2,563	0.149	0.357	0	1
Agri. Cooperative	2,563	0.0222	0.147	0	1
Agri. Committee	2,563	0.139	0.346	0	1

Socio-economic characteristics. The average community size is 64 households (**Table 1**).⁴⁵ The households head is predominantly male. Around half of household heads are literate and 30% completed some or all of primary school. Road infrastructure in the region is underdeveloped, causing transport costs

⁴⁵ By construction this is below the regional community size of about 200 households as we restricted our sample to communities with fewer than 100 households (Bulte et al., 2015)

to heavily influence farmer decision making regarding purchasing of market inputs and sale of crops produced. Within our sample, the average cost of transport to reach the closest market is 1,780 Cf, roughly equivalent to USD\$1.83. Disputes surrounding access use and management of land in the Kivu provinces are complex and are closely linked with the root causes of conflict within the region (Huggins, 2010). We find that communities within our sample experience an average of one land dispute annually, although variation between communities is relatively high. In the absence of a functional and effective central governing body present within the South Kivu province, local chiefs are largely responsible for governing daily life in their communities through customary law. Chief status is hereditary and held for life, with chiefs in our sample having been in the position for an average of 26 years. People interact with the chief during about 6 days of the month and roughly 30% of household heads hold some formal position of leadership within the community. Agricultural group membership is low, with labor groups and committees being the most popular and nearly zero participation in cooperatives. This near absence of cooperative membership highlights the absence of formal support for marketing and sales of crop produce for farmers within the region.

Selecting Contact-farmers. Our main survey variable below, is a hypothetical question in which we asked household heads to nominate one individual in the community who they felt is best suited for spreading information about new agricultural technologies. We follow the approach in Banerjee et al. (2013), albeit adapted to the agricultural context, and ask:

If we wanted to spread information about a new agricultural technique, to whom do you suggest we speak in this community?

Community members are intrinsically motivated by two factors in making their nomination: (i) the perceived reliability of the nominee in learning about and distributing resource and information, and (ii) the likelihood of personally benefiting indirectly through the nominee. We explicitly frame the question as an information distribution question and not selection of an extension contact-farmer. This is to minimize any ulterior motivations behind nomination of social hierarchy or social control by which resource or financial benefits could be personally captured. The candidacy question is therefore framed to capture innovation adoption through improved information, emphasizing influence and reach of nominees.

On average within each community, 90% of household heads provided a nomination in response to the candidacy question (**Table 2**) and 12 different household heads were put forward (21% of all community households). From these nominations we select from each community the three candidates with the highest number of nominations. The average number of nominations received by candidate household heads is 2.6 (1% of community nominations), but this average increases to 8.37 for top nominees (16%).⁴⁶

Table 2 Summary Statistics (Level and Percentage) of All Candidate Nominations, Top Candidate Nominations, and Nominators within Communities

VARIABLES	Level				Percentage of Community			
	Mean	s.d.	Min	Max	Mean	s.d.	Min	Max
<i>All Candidates</i>								
# per Community	13.95	4.39	4	23	21%	0.060	10%	39%
# Nominations per Candidate	2.59	1.09	1	6.75	1%	0.006	0%	3%
<i>Top Candidates</i>								
# per Community	3	0	3	3	5%	0.02	4%	13%
# Nominations per Candidate	8.47	4.28	1	17.5	16%	0.080	5%	38%
<i>Nominators</i>								
# per Community	55.25	14.55	21	73	90%	0.100	57%	100%

“Level” values are not normalized by community size while “Percentage of Community” normalizes values by total number of households within the community. Number of top candidates in Level values is equal to 3 by study design and does not vary across communities but percentage values will vary due to differences in community sizes.

Figure 1 plots the frequency of nominations occurring by shortest path length. Nominations are most frequently made along direct network links and drop off rapidly with each subsequent path length. Incidence of self-nomination is very low.

Figure 2 separates out nominations by relationship type with nominations made through indirect links grouped in a separate column. Relationship ties are not mutually exclusive so multiple relationships along a single network link do occur. Based on **Figure 2**, no relationship tie dominates the others within nomination decisions.

⁴⁶ As our candidacy question did not restrict nominations to only household heads, there are individuals who were nominated that are not included within our sample. Thus it is not a direct relationship that all 90% of household heads nominations are reflected within the FT and top FT metrics reported in Table 4.

Figure 1 Path Length Frequencies of Nomination

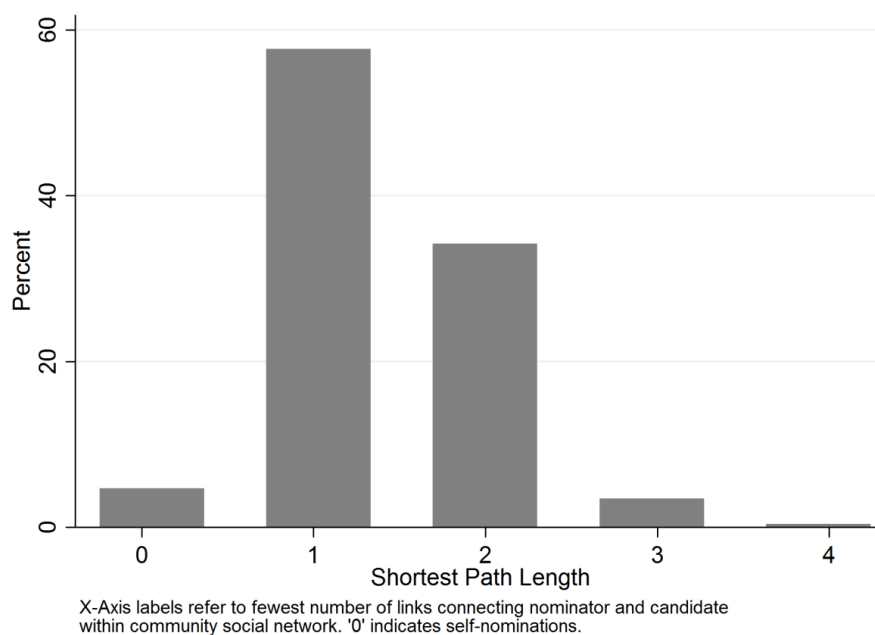
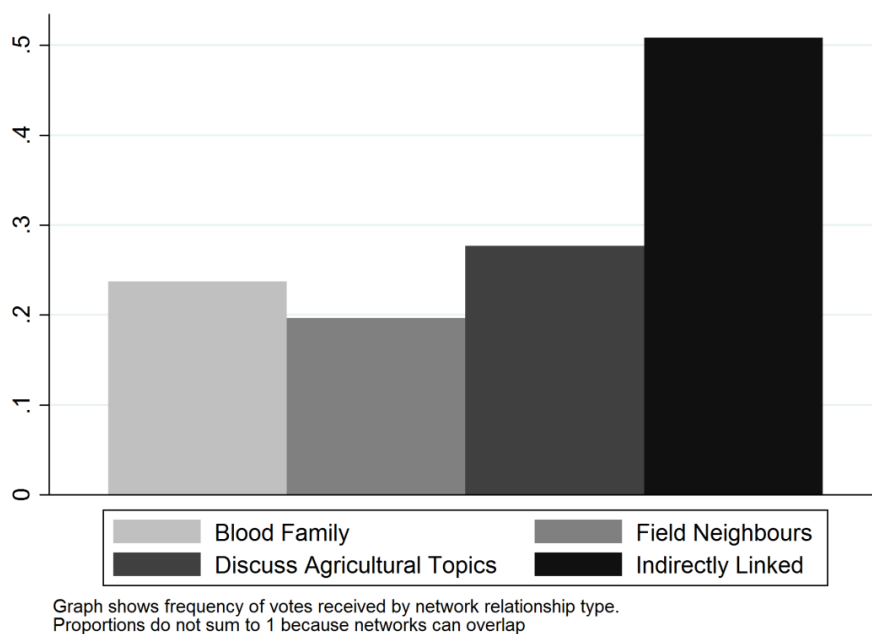


Figure 2 Frequencies of Nominations by Relationship Type



Social networks. In mapping social networks, we focus on three key relationships of interest: family, field neighbors and people that often discuss agricultural issues together. Each of these relationships form the basis not only for frequent interaction, and thus direct knowledge of a candidate’s merits, but also channels of information and resource distribution through which an individual may ultimately benefit within an

agricultural context.⁴⁷ In collecting network data along each of the three social dimensions concerns of measurement error resulting from memory bias was a chief concern. To overcome this risk, each respondent was read the full roster of community members compiled through the community census repeated for each social dimension. Respondents would respond “yes” or “no” to each name read indicating the presence of a connection along the social dimension in question.⁴⁸

In total, we capture information on 2,563 individuals and 20,596 relationships. Social ties are then aggregated into a single combined network graph for each community. The resulting dataset is thus an unweighted directional network. The network structures of communities exhibit exceptionally clustered network structures and short diameters relative to network size, signifying quite dense network structures.

Figure 3 *Sample Community Social Network Structure*

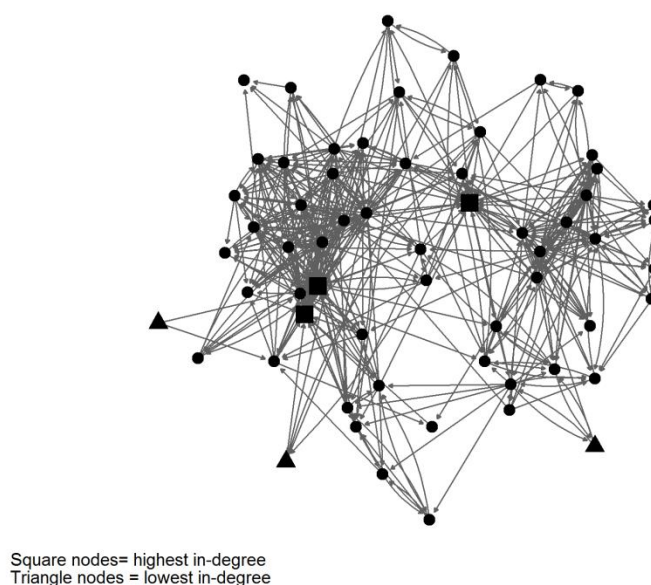


Figure 3 graphically presents a sample network structure for a community within our sample. The most connected household heads, measured by in-degree, are colored green while the least connected household heads are colored red. There is strong variation evident in connectedness levels of the households. This variation in network connectedness of households is representative of the larger sample.

⁴⁷ Early pilot tests included multiple other social dimensions within the network survey but found no added value in variance captured with their inclusion. To minimize potential respondent survey fatigue, we restricted network data to these three dimensions.

⁴⁸ Household heads and respective wives were independently named on the community census while all other adult household members were grouped under a label of “other household member”.

Social preferences. To elicit proxies of an individual’s social preferences, we use four hypothetical allocation decisions based on Fehr et al (2008) in which respondents are asked to choose between two sets of payoff scenarios for themselves and an anonymous member of the community (see Table A1 for choice options). We classify participants as selfish (who maximize their own pay-off), generous (those that maximize receiver pay-off) and inequality averse. The rest is categorized as other (the omitted category for our analyses). The majority of respondents (60%) are inequality averse.

We also include a hypothetical dictator game question to elicit respondent sharing behavior of an allotment of 3 kg of chemical fertilizer. They are able to choose with whom they would like to share as well as the amount they would like to give away. On average respondents are willing to share just under half of the endowment.

4 Empirical Analysis

Our study addresses two primary research questions. First we ask, what characteristics drive community nominations of contact-farmers? We include within our analysis traditional socioeconomic variables as well as measures of an individual’s social preferences and network position. In addition, we zoom in on the direct relationships between nominators and the household heads they nominate to understand how relative differences in characteristics determine nominations. Second, we explore the extent that community nominations overlap with households identified as maximum diffusion entry-points in the complex contagion simulations. This comparison allows us to assess whether contact-farmer selection by community members sufficiently aggregates each individual’s incomplete information set on networks for harnessing network diffusion dynamics.

4.1 *Characteristics of Community Nominated Contact-farmers*

We set out by exploring what characteristics are most influential for community nominations of contact-farmers. Specifically, we estimate:

$$Y_{iz} = \alpha_0 + \alpha_1 S_{iz} + \alpha_2 N_{iz} + \alpha_3 X_{iz} + \alpha_4 C_z + \varepsilon \quad (1)$$

where, Y_{ij} is an indicator for whether household head i (with, $i = 1, \dots, H_j$ where H_j is the number of households in community z) of community z (with $z = 1, \dots, 40$) received any nomination. S_{iz} is a vector containing our social preference measures, N_{iz} contains individual network position indicators (eigenvector centrality measures and clustering coefficient), X_{iz} is a vector of socioeconomic and agriculture covariates, C_z is a vector of community fixed effects, and ε is an error term.

To fully capture what characteristics were most correlated with nominations we re-run equation (1) and substitute the dependent variable to capture whether household head i is among the three most nominated candidates in a community (termed Top Candidates) as well as the proportion of nominations household head i received relative to total nominations made within community z , which provides a normalized measure of vote-capture.

Next we focus on nomination decisions between nominators and candidates. We do so by exploring the relational differences in characteristics between pairings where a nomination occurred versus pairings where no nomination occurred. We first construct a variable capturing the difference between each pair of household heads within community z . We denote this differenced variable with a bar, in which i is a household head potentially nominated by household j .

$$\Delta c_{ij} = c_i - c_j$$

Difference variables are not taken in absolute value form so that direction within the differences are maintained. All socioeconomic, social preference, and network position indicators within our dataset are differenced for every combination of household pairs within community z . We then estimate the dyadic model:

$$Y_{ijz} = \alpha_0 + \alpha_1 \Delta S_{ijz} + \alpha_2 \Delta N_{ijz} + \alpha_3 \Delta X_{ijz} + \alpha_4 C_z + \varepsilon_{ij} \quad (2)$$

where, Y_{ij} is an indicator for whether household head i nominated household head j (with $j = 1, \dots, H_z$ and $i \neq j$) within community z . Term ΔS_{ijz} , is a vector containing the differenced social preference measures for i and j , ΔN_{ijz} contains the difference in individual network position scores (eigenvector centrality measures and clustering coefficient), ΔX_{ijz} is a vector of differenced socioeconomic and agriculture covariates. C_z is

a vector of community fixed effects, and ε_{ij} is an error term clustered on both household i and j within community z .⁴⁹

4.2 Comparing Complex Contagion-Selection with Community Nomination

We use threshold complex contagion model simulations to identify the two individuals best suited to be injection points into the community network.⁵⁰ We follow Beaman et al (2015) and first randomly assign each individual in the network a normally distributed threshold, with mean of 2 and standard deviation of 1. Then, two household heads in the network are selected at random to be the network entry point and are deemed ‘infected’.⁵¹ Any individuals that are connected to at least as many of the infected nodes as their assigned threshold are then also considered infected.

The simulation is continued for four iterations.⁵² Within each iteration, any household who exceed their threshold in links to infected individuals also become infected. We do not allow for households to become un-infected within the simulation. After the four possible infection periods, the proportion of household heads that are infected is calculated. This proportion measures the simulation diffusion rate. For each possible pair of initial targets within the community, we repeat this same simulation 2000 times to account for the element of the randomness built into the model. The two individuals that achieve, on average, the highest diffusion rate within the complex contagion selection process are selected as the simulation candidates.

Next, we compare the relationship between community polling candidates and simulation candidates. We use a logit model where the dependent variable, Y_{iz} is an indicator for household i receiving in the top 3 highest number of community nominations. Specifically we estimate:

$$Y_{iz} = \beta_0 + \beta_1 \text{Simulation}_{iz} + \beta_2 C_z + \varepsilon \quad (3)$$

Where Simulation_{iz} takes the value 1 if household i is a simulation candidate and 0 otherwise. C_z is a vector of community fixed effects and ε is the error term. We re-run Equation (3) on the proportion of

⁴⁹ This is the suggested method for clustering standard errors within regressions of dyadic relationships (see Cameron, Gelbach, & Miller, 2011; Petersen, 2009; Thompson, 2011)

⁵⁰ We do this for pairs and not triplets due to the computational complexities render triplets time prohibitive: the current model required a running time of one day for completion while triplets require a total running time of nearly 500 days.

⁵¹ Where ‘infection’ can be considered as having received information or having adopted the innovation based on the specifics of what diffusion process is being modeled.

⁵² In models with no reverting from adopter to non-adopter and infinite time periods the model will convergence on perfect adoption. Therefore we set a time period limit of 4 iterations.

nominations household i received within community \mathcal{Z} as well as a comparison through two restricted version of the full community polling process.⁵³

These two alternative sampling of processes are often observed within farmer-led extension programs in practice (Beaman et al., 2015; Simpson et al., 2015; Bulte et al., 2015), (i) we restrict our nominees to those selected by community chiefs and other leaders within the community⁵⁴ and (ii) select farmers based on eligibility criteria commonly used by NGOs (see Bulte et al, 2015) which include farmers that own land, are farming as a primary activity, are literate, have over 10 years farming experience, and have previously used in-organic fertilizer.⁵⁵ In essence, both alternative selection process are restricted versions of the full community polling process.

5 Results

5.1 *Community Selection of Farmer Trainers*

We set out by assessing the characteristics of nominated candidates as selected by community members. Table 3 reports the marginal effects (at the mean). In Column 1, we assess characteristics of individuals who received at least one nomination in the candidacy questions. Column 2 restricts out dependent variable to three most frequently mentioned nominees and Column 3 assesses the proportion of nominations an individual received.

Across the columns we observe similar patterns. Focusing first on our social network and preferences variables, we see that nominated community members have higher centrality scores and are more likely to behave more pro-socially in a dictator game setting. These results corroborate research on social influence (Walsh et al., 2004). Despite Centola (2010) findings of clustering measures affecting diffusion, our results

⁵³ All alternative selection processes were conducted ex-post on our community-wide polling dataset and not conducted directly within communities.

⁵⁴ Leaders included are the chief, under-chiefs, and the chief's advisors ("Mushamuka")

⁵⁵ Use of fertilizer greatly restricts the sample size, resulting in some communities have less than 2 individuals meeting all criteria. In these instances, those individuals meeting all criteria are ranked first with individuals meeting all but the fertilizer use criteria included in order of nominations received.

Table 3 Regression Results on dependent variables nomination as a Farmer-Trainer (FT), top-nominated FT, and proportion of nominations

	(1) Candidate	(2) Top Candidate	(3) Prop. of Nominations
Model	Logit	Logit	Tobit
Network Position			
Eigenvector Centrality	0.146*** (0.031)	0.062*** (0.015)	0.065*** (0.011)
Clustering Coefficient	0.026 (0.038)	-0.041* (0.023)	0.003 (0.013)
Social Preference Indicators			
Selfish	-0.010 (0.036)	-0.012 (0.018)	-0.007 (0.013)
Generous	0.036 (0.030)	0.009 (0.013)	0.013 (0.011)
Inequality Averse	0.014 (0.019)	-0.003 (0.009)	0.004 (0.007)
Amount shared	0.053*** (0.020)	0.007 (0.009)	0.015** (0.007)
Household Characteristics			
Native Language	0.014 (0.069)	0.019 (0.049)	0.005 (0.024)
Female Head (=1)	-0.102*** (0.022)	-0.052*** (0.020)	-0.035*** (0.008)
Age	0.013*** (0.003)	0.003** (0.002)	0.004*** (0.001)
Age^2	-0.000*** (0.000)	-0.000* (0.000)	-0.000*** (0.000)
Household size	0.008*** (0.003)	0.002* (0.001)	0.003*** (0.001)
Migrant (=1)	0.027 (0.019)	0.011 (0.009)	0.010 (0.007)
<i>Highest Education Attained</i>			
Primary	0.066*** (0.020)	0.034*** (0.011)	0.028*** (0.007)
Secondary	0.137*** (0.022)	0.036*** (0.012)	0.046*** (0.008)
Tertiary	0.177*** (0.055)	0.023 (0.022)	0.047** (0.019)
War exposure score	0.003 (0.009)	-0.003 (0.004)	-0.002 (0.003)
Frequency Talks Chief	0.000 (0.001)	0.001** (0.000)	0.001* (0.000)
Wealth Index	0.026*** (0.008)	0.007* (0.004)	0.009*** (0.003)
Community Leader	0.095*** (0.017)	0.025*** (0.009)	0.035*** (0.006)
Agricultural Indicators			
Years Farming	0.000 (0.001)	0.000 (0.001)	-0.000 (0.000)
Primary Farmer	-0.001 (0.020)	0.005 (0.010)	0.002 (0.007)
Fertilizer Knowledge	-0.002 (0.006)	-0.000 (0.003)	-0.001 (0.002)
Agri. Labor Group	0.010 (0.022)	0.008 (0.010)	0.006 (0.008)
Agri. Cooperative	-0.019 (0.052)	0.024 (0.017)	-0.009 (0.018)
Agric. Committee	0.035 (0.022)	0.020** (0.009)	0.017** (0.008)
Observations			
N	2,335	2,201	2,335

find limited effect of clustering on nomination and even a weakly negative relationship for Top Candidates. It is possible that the centrality of an individual is more accurately approximated through observation relative to clustering. While the first can be approximated rather crudely based on how connected an individual is to other well connected community members, the second is more challenging as it requires more explicit knowledge of a greater number of social links within the network structure.

Several socioeconomic characteristics correlate with community nomination. Being male, older, head of larger household, educated, in a community leadership position, and of greater wealth are all positively correlated with both the probability of nomination as well as the magnitude of nominations received by an individual. These correlations generally reflect characteristics of prominent individuals within a community. We find little evidence of topical expertise being positively correlated with probability of nomination or proportion of nominations received. Only membership on a community agricultural committee is found to be correlated within our agricultural indicators.

We zoom further in on the presence of homophily within nomination decisions. **Table 4** presents log odds-ratios of the difference in indicators between households on an indicator of a nomination occurring.

We observe similar patterns in the direct relationships between households and their nominated candidates that were observed at the community level in **Table 3**. For the network and social preference indicators, we observe that households are more likely to nominate candidates that are more central, less clustered, and more pro-social relative to themselves. In household indicators, nominated candidates are more likely to be male, head of a larger household, have experienced less war exposure, talk more frequently with the chief, wealthier, and better educated than their nominators. Where we see a deviation of individual nomination correlates from the group-level correlates is in agricultural indicators. Nominators tend to nominate individuals who have *lower* fertilizer knowledge and are less likely to be members of agricultural cooperatives relative to themselves. This reinforces the above finding that less emphasis is placed on domain expertise and more on perceived social influence within nomination decision processes.

Table 4 Regression Results on Homophily within Nominations

Differenced Variables ($j - i$)	(1) Nomination Cast
Network Indicators	
Δ Eigenvector Centrality	0.871*** (0.165)
Δ Clustering Coefficient	-0.243 (0.190)
Social Preference Indicators	
Δ Generous	-0.191 (0.145)
Δ Selfish	-0.271* (0.162)
Δ Inequality Averse	-0.121 (0.146)
Δ Amount Shared	0.108 (0.101)
Household Indicators	
Δ Female	-0.179** (0.087)
Δ Native Language	1.017** (0.516)
Δ Age	0.007 (0.008)
Δ HH Size	0.054*** (0.018)
Δ Migrant	0.179 (0.109)
Δ War Exposure	-0.117*** (0.034)
Δ Community Leader	0.421*** (0.115)
Δ Frequency Talks Chief	0.023*** (0.009)
Δ Wealth Index	0.238*** (0.084)
Δ Education Level Attained	0.419*** (0.116)
Agricultural Indicators	
Δ Farming Experience (yrs.)	0.002 (0.008)
Δ Primary Farmer	0.008 (0.085)
Δ Fertilizer Knowledge	-0.065* (0.035)
<i>Agricultural Group Membership</i>	
Δ Labour Collective	0.242 (0.167)
Δ Cooperative	-0.556** (0.275)
Δ Community Committee	0.268 (0.187)
Constant	-5.555*** (0.440)
Observations	145,727

*** p<0.01, ** p<0.05, * p<0.1

Village fixed effects included. Logistic regression, coefficients report odds-ratio estimates. Differenced Variables are (j = Potential Candidate) - (i = Potential Nominator) for all possible pair combinations within each village. Standard errors clustered on two levels: potential candidate (j) and potential nominator (i)

5.2 Simulation-selected Comparison

Our exploratory results indicate that community members are accounting for network position indicators in addition to pro-social preferences and socio-economic qualities in selecting contact-farmers. However, it is unclear whether those selections are optimal targets as would be predicted from contagion simulation models. We graphically compare this overlap for full community polling, two restricted-polling processes, and random selection processes in **Figure 4**.

Figure 4 Alternative selection processes and their overlap with optimal complex contagion entry points

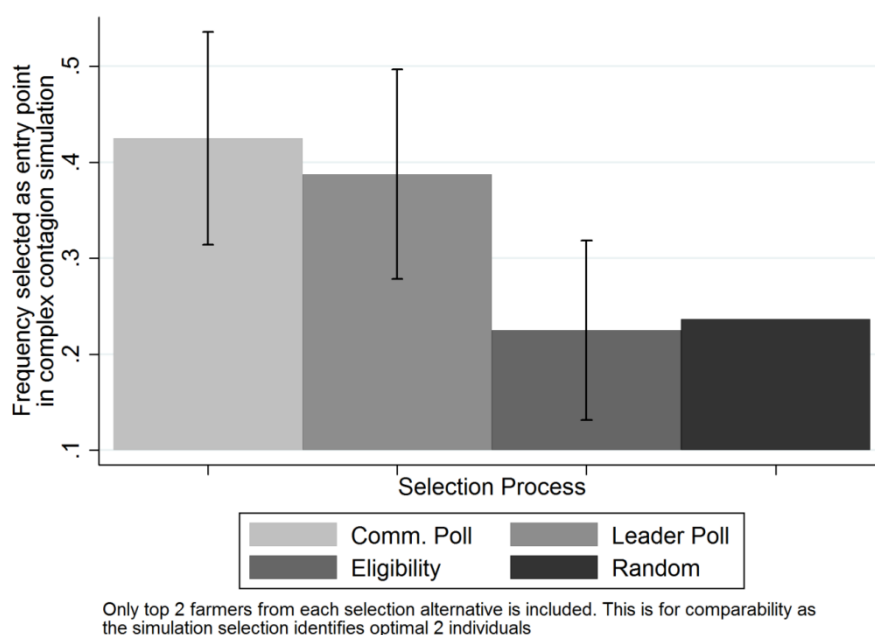


Figure 4 shows that random selection of entry-points results in optimal entry-points being selected only 25% of the time. However selection by community-wide polling identifies optimal entry-points in over 40% of instances. This result indicates that community members are incorporating factors that relate to diffusion and network-influence within their nomination decisions. Restricting nominations to only community leaders has only a small impact on selection of optimal entry-points relative to full community nomination. However restriction of who can be a candidate, based on eligibility criteria, substantially decreases the probability of optimal entry-points being selected. The imposition of eligibility-criteria reduces probability of selecting optimal entry-points down to the same level as random selection, erasing any gains that community-wide nomination may generate. This result indicates that restricting the pool of nominators

to only community leaders has considerable less impact on resulting selection of optimal entry-points than restricting the pool of candidates who can be nominated.

Turning only to the three community nomination-based approaches, we see that full community-polling outperforms the two restricted community-polling alternatives.. The difference in means is not significant relative to leader-restricted polling but is for restricting candidacy based on eligibility criteria. **Table 5** presents formal empirical results, conditional on community fixed effects.⁵⁶ Column (1) reports the marginal effects (at the mean) of being identified as a candidate through simulation on being nominated as a Top Candidate in the community wide polling process. Column (2) looks at the same relationship on the proportion of nominations received. Columns (3) and (4) analyze the two restricted selection processes of leadership polling and eligibility-criteria for candidates.

Table 5 Regression Results on Overlap Complex Contagion Optimal Entry Points with Candidates, Top Candidates, and Proportion of Nominations Received

VARIABLES	(1) Top Candidate	(2) Prop. of Nominations	(4) Leader Candidate	(5) Eligibility Candidate
Simulation Candidate	0.399*** (0.056)	0.273*** (0.019)	0.353*** (0.054)	0.174*** (0.044)
Constant		-0.122*** (0.027)		
Observations	2,563	2,537	2,563	2,563

*** p<0.01, ** p<0.05, * p<0.1

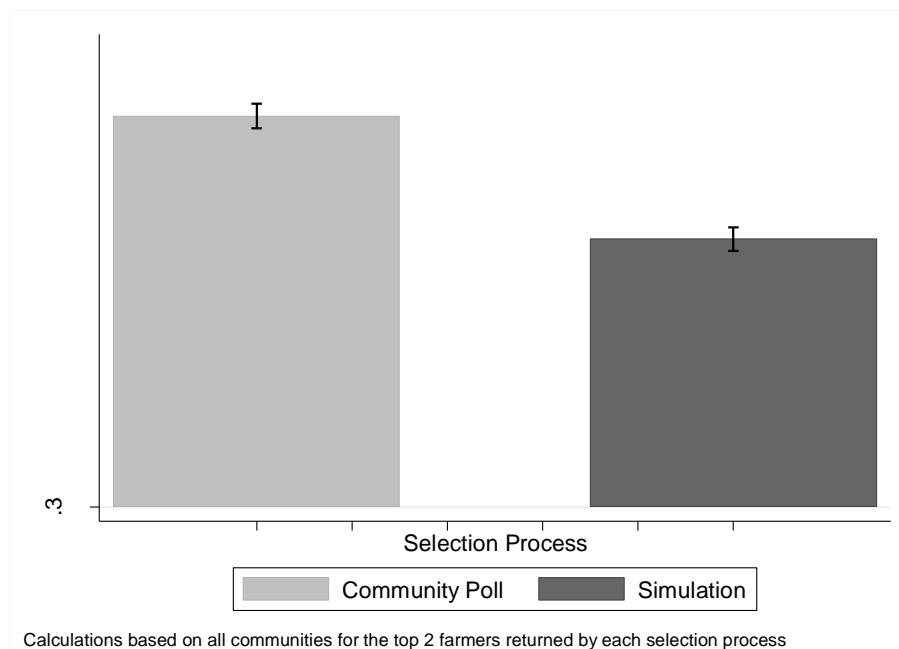
Logistic Regressions, coefficients report marginal effect estimates. Village fixed effects included. Complex Contagion (=1) indicates individual i was selected as one of the optimal entry points in a complex contagion simulation. Standard errors clustered at community level in parentheses

All three of the community-based selection processes and the proportion of nominations received are found to be positively correlated with selection as simulation-identified candidates. These results indicate that communities are implicitly nominating those individuals who are more likely to be identified as optimal entry-points using complex-contagion targeting (as seen in Beaman et al., 2015). However community selection and network-based selection is not perfectly correlated. This imperfect correlation is only problematic if it results in substantial losses in diffusion rates. **Figure 5** presents a graphical comparison of the diffusion rates achieved within simulations where the simulation candidates are used as entry points with

⁵⁶ We adjust our definition of top farmer from top three to top two so that sample sizes of community polling selection processes and complex contagion selected individuals are equivalent.

those rates achieved by community-nominated Top Candidates.⁵⁷

Figure 5 Estimated Diffusion Rates for Complex Contagion Optimal Entry Points (Simulation) and Top Community Nominated Candidates (Comm. Poll)



A difference in diffusion rates of less than 10% is observed, although this difference is statistically significant. In our sample of communities, this diffusion rate difference would translate into an average of 6 fewer households being reached when community Top Candidates are used as entry points instead of simulation candidates. Whether this trade-off in reach for cost-savings is worthwhile for a development practitioner will depend on factors of financial endowments, network data access costs, innovation being spread, and program objectives.

6 Conclusion

Adoption of innovations is of central importance for development and is particularly critical for increasing agricultural productivity within under-developed regions. Determining optimal entry points into social networks has been identified as a key concern in achieving maximum spread of information and innovations. Network-theory based selection has been found to improve the width of diffusion of agricultural

⁵⁷ Only full community-polling is compared as it was the top-performing selection process and will give an indication of how a difference in candidate-selection overlap translates to a difference in diffusion rates.

innovations, but requires collection of costly network data (Beaman et al., 2015). Justifying these high costs requires evidence of sufficient returns in diffusion, and ultimately program impacts. This study speaks to the first step of this trade-off. We use detailed network data and low-cost community polling to compare community-nominated entry points against simulated optimal entry points. Specifically, we contextualize the polling question to agricultural information and set diffusion simulation parameters to reflect that adoption of improved agricultural practices more closely approximate complex-diffusion processes (Beaman et al., 2015). We explore first who is being nominated by communities combining the community poll results with detailed network, social preferences, socio-economic, and agricultural indicators. We then explore how closely community selection of candidate entry points correlates with simulation-identified optimal entry points.

At both the aggregate level and within individual nomination decisions similar patterns emerge around which community members are being nominated. Selected candidates reflect cultural indicators of ‘success’, being male, educated, older, wealthier, and more involved in community leadership. Additionally, candidates are more likely to be network central and to be more pro-social. Domain expertise has been found in previous studies to be a factor in one’s level of social influence for specific innovations (Chan & Misra, 2013; Leonard-Barton, 1985). However, we do not see any obvious signs of communities incorporating agricultural indicators within their nomination decisions. This indicates that communities may be nominating based on perceptions of more general social prominence. However this result could be explained by the framing of the candidacy question, as the question states “*agricultural information*” which offers little indication on whether expertise is required for understanding the information. Furthermore framing the question using “*agricultural technology*” may also yield different results.

In comparing community selected candidates to those identified through complex contagion simulations as being optimal, we see that the two sets overlap in 40% of cases. Even reducing the polling sample down from full community to only leaders within the community yields similar results. In looking at resulting differences in simulated diffusion rates for simulation optimal against community selected candidates, yields a drop in diffusion of less than 10 percentage points. These results indicate that community members are, at the very least, implicitly aware of diffusion dynamics and could prove a valuable resource in identifying network entry points. Where network data is costly to collect, development practitioners aiming to spread

new innovations have the alternative of using community polling in selecting their contact points.

One result that stands out as being critical for program design is that the eligibility-criteria approach to selecting community contact-points is potentially detrimental to overall diffusion. Such eligibility-based approaches are common within agricultural extension (Anderson & Feder, 2007; Beaman, Udry, Karlan, & Thuysbaert, 2013; Simpson et al., 2015). Our results indicate that development practitioners would be better to using community-nomination selection over eligibility-selection, an alternative approach with greater impacts on diffusion with the added benefit of being lower cost to implement.

This study only compares diffusion rates using simulation and imposes exogenous parameters on the likelihood of a given individual to adopt. Real world diffusion is considerably more complicated than our simulation model. While evidence is emerging on the predictive power of simulations for real world outcomes in a developing countries, it remains very limited (Beaman et al., 2015). Quantifying the true trade-off between data-demands and diffusion rates requires further testing under field experiment conditions until such predictive power is found to be robust. Given the costliness of field-experiments, more confidence must first be had in identifying *which* levers must be tested (e.g. the network parameters used for selection; the framing of the candidacy question, etc). Network simulations offer a cost-effective means of identifying and prioritizing such levers using less cost-intensive observational data.

Finally, our study was conducted in small communities of rural eastern Democratic Republic of Congo where networks are dense and highly connected. As information regarding underlying network structures is inferred through observation of social interactions, scaling up to larger communities may have different results. Further research into the relationship between community size and member awareness of underlying network structures is necessary to determine how well our results translate to large communities.

References

- Acemoglu, D., & Ozdaglar, A. (2011). Opinion Dynamics and Learning in Social Networks Opinion Dynamics and Learning in Social Networks. *Dynamic Games and Applications*, 1(1), 3–49.
- Anderson, J. R., & Feder, G. (2004). Agricultural extension: Good intentions and hard realities. *World Bank Research Observer*, 19(1), 41–60. <https://doi.org/10.1093/wbro/lkh013>
- Anderson, J. R., & Feder, G. (2007). Chapter 44 Agricultural Extension. *Handbook of Agricultural Economics*, 3(6), 2343–2378. [https://doi.org/10.1016/S1574-0072\(06\)03044-1](https://doi.org/10.1016/S1574-0072(06)03044-1)
- Aral, S., & Walker, D. (2012). Identifying Influential and Susceptible Members of Social Networks. *Science*, 337, 337–341. <https://doi.org/10.1126/science.1215842>
- Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2014). *Gossip: Identifying Central Individuals in a Social Network* (NBER Working Paper Series No. 20422). *National Bureau of Economic Research*. <https://doi.org/10.3386/w20422>
- Beaman, L. a, Udry, C. R., Karlan, D., & Thuysbaert, B. (2013). Profitability of Fertiliser: Experimental Evidence from Female Rice Farmers in Mali. *The American Economic Review*, 103(3), 381–386. <https://doi.org/10.1257/aer.103.3.381>
- Beaman, L., Benyishay, A., William, C., & Mobarak, A. M. (2015). *Can Network Theory based Targeting Increase Technology Adoption ?*
- BenYishay, A., & Mobarak, A. M. (2014). Social Learning and Communication. *NBER Working Paper*, (20139). <https://doi.org/10.3386/w20139>
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27, 55–71. <https://doi.org/10.1016/j.socnet.2004.11.008>
- Bulte, E., Leuvel, K., Nillesen, E. E. M., Pieters, J., & Voors, M. (2015). *Farm Households in Eastern Congo, Baseline Survey Report*.
- Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust Inference With Multiway Clustering. *Journal of Business & Economic Statistics*, 29(2), 238–249. <https://doi.org/10.1198/jbes.2010.07136>
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329, 1194–1197. <https://doi.org/10.1126/science.1185231>
- Centola, D. M., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3), 702–734. <https://doi.org/10.1086/521848>
- Chan, K. K., & Misra, S. (2013). Characteristics of the Opinion Leader: A New Dimension. *Journal of Advertising*, 19(3), 53–60. <https://doi.org/Article>
- Cho, Y., Hwang, J., & Lee, D. (2012). Identification of effective opinion leaders in the diffusion of technological innovation: a social network approach. *Technological Forecasting and Social Change*, 29(1), 97–106.
- Coghlan, B., Ngoy, P., Mulumba, F., Hardy, C., Bemo, V. N., Stewart, T., ... Brennan, R. (2007). *Mortality in the Democratic Republic of Congo: An Ongoing Crisis*. *International Rescue Committee*. Retrieved from http://www.rescue.org/sites/default/files/migrated/resources/2007/2006-7_congomortalitysurvey.pdf
- Davis, R. L. (1999). Using Opinion. *Annals of the American Academy of Political and Social Science*, 566, 55–67.
- FAO. (2015). *FAO Statistical Pocketbook 2015. Food and Agriculture Organization of the United Nations*. <https://doi.org/978-92-5-108802-9>
- Feder, G., Anderson, J. R., Birner, R., & Deininger, K. (2010). Promises and Realities of Community-Based Agricultural Extension. *Community, Market and State in Development*, (March), 187–210. https://doi.org/10.1057/9780230295018_12
- Feder, G., Murgai, R., & Quizon, J. B. (2003). *Sending Farmers Back to School: The Impact of Farmer Field Schools in Indonesia* (No. 3022). *Policy Research Working Paper*. Retrieved from <http://econ.worldbank.org>
- Feder, G., & Savastano, S. (2006). *The role of opinion leaders in the diffusion of new knowledge: the case of integrated pest management* (World Bank Policy Research Working Paper No. WPS3916).
- Granovetter, M. (1978). Threshold Models of Collective Behavior. *Threshold Models of Collective Behavior*, 83(6), 1420–1443.
- Huggins, C. (2010). *Land, Power, and Identity: Roots of violent conflict in Eastern DRC*.
- Kendzior, J., Zibika, J. P., & Voors, M. (2015). *Social relationships, local institutions, and the diffusion of improved variety seed and field management techniques in rural communities: six case studies in South Kivu, DRC*.
- Kondylis, F., Mueller, V., & Zhu, S. J. (2014). *Seeing is believing? Evidence from an extension network experiment* (Policy Research Working Paper No. 7000).
- Leonard-Barton, D. (1985). Experts as Negative Opinion Leaders in the Diffusion of a Technological Innovation. *Journal of Consumer Research*, 11(4), 914. <https://doi.org/10.1086/209026>
- Lukuyu, B., Place, F., Franzel, S., & Kiptot, E. (2012). Disseminating Improved Practices: Are Volunteer Farmer Trainers Effective? *The Journal of Agricultural Education and Extension*, 18(5), 525–540. <https://doi.org/10.1080/1389224X.2012.707066>
- Pentland, A. (2014). *Social Physics: How good ideas spread-the lessons from a new science*. Penguin.
- Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22(1), 435–480. <https://doi.org/10.1093/rfs/hhn053>
- Pypers, P., Sanginga, J. M., Kasereka, B., Walangululu, M., & Vanlauwe, B. (2011). Increased productivity through integrated soil fertility management in cassava-legume intercropping systems in the highlands of Sud-Kivu, DR Congo. *Field Crops Research*, 120(1), 76–85. <https://doi.org/10.1016/j.fcr.2010.09.004>
- Simpson, B. M., Franzel, S., Degrande, A., Kundhlande, G., & Tsafack, S. (2015). Farmer-to-Farmer Extension: Issues in Planning and Implementation. *MEAS Technical Note*, (May).
- Sperling, L., & Loevinsohn, M. E. (1993). The dynamics of adoption: Distribution and mortality of bean varieties among small farmers in Rwanda. *Agricultural Systems*, 41(4), 441–453. [https://doi.org/10.1016/0308-521X\(93\)90044-3](https://doi.org/10.1016/0308-521X(93)90044-3)

- Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, 99(1), 1–10. <https://doi.org/10.1016/j.jfineco.2010.08.016>
- Valente, T. W. (1996). Network models of the diffusion of innovations. *Computational & Mathematical Organization Theory*, 2(2), 163–164.
- Valente, T. W., & Pumpuang, P. (2007). Identifying Opinion Leaders to Promote Behavior Change. *Health Education & Behavior*, 34(6), 881–896. <https://doi.org/10.1177/1090198106297855>
- Van Eck, P. S., Jager, W., & Leeftang, P. S. H. (2011). Opinion leaders' role in innovation diffusion: A simulation study. *Journal of Product Innovation Management*, 28(2), 187–203. <https://doi.org/10.1111/j.1540-5885.2011.00791.x>
- Vlassenroot, K., & Raeymaekers, T. (2004). *Conflict and social transformation in Eastern DR Congo*. Academia Press.
- Walsh, G., Gwinner, K. P., & Swanson, S. R. (2004). What makes mavens tick? Exploring the motives of market mavens' initiation of information diffusion. *Journal of Consumer Marketing*, 21(2), 109–122. <https://doi.org/10.1108/07363760410525678>
- WORLD, B. (2015). World Bank Development Indicators. Retrieved from data.worldbank.org/indicator/NY.GDP.PCAP.CD

Appendix

Table A1 Variable names, descriptions and constructed minimums and maximums values

Variable Name	Variable Description	Min.	Max.
Household (HH) Size	Number of household members	1	20
Age	Age in Years of Household Head	14	103
Female	Household Head is female (=1)	0	1
Married	Household head is married (=1)	0	1
Migrant	Household head was not born in the community (=1)	0	1
Literate	Household head is literate (=1)	0	1
Wealth Index	Normalized wealth index constructed using a weighted mean effects index on variables of livestock owned, farm size, and land tenure. Note that outlying farm sizes in excess of 100 sq. km. were dropped from the sample.	-2.2	12.9
Talks Chief	Number of days household head talked to the community chief in the last month	0	31
War Exposure	War exposure score calculated as Saw Fighting, 0.5 point. Property Damage/Loss, 0.5 points. Injured in the war, 1 point. Family member injured in the war, 0.5 point. Family member killed in the war, 1 point. Migrated because of the war, 1 point.	0	4.5
Native Language	Household head is a native speaker of the local language (=1)	0	1
Leadership Position	Household head holds a leadership position within the community (community chief excluded). Possible positions include quarter chiefs, religious leaders, educational leaders, health leaders, youth leaders, women leaders, and agricultural group leaders.	0	1
Primary	Household head's highest educational attainment was some or all of primary school.	0	1
Secondary	Household head's highest educational attainment was some or all of secondary school.	0	1
Tertiary	Household head's highest educational attainment was some or all of a tertiary school.	0	1
Fertilizer Knowledge	Fertilizer knowledge score calculated using correct answers to questions: Specifically, the index is based on the following questions: <i>What is the effect of fertilizer on yields? What are other effects of fertilizer? When is it effective to apply fertilizer? What is the best method to apply fertilizer? What is the price of fertilizer in Bukavu?</i> Maximum possible score of 7.5	0	6.5
Years Farming	Number of years the household head has been farming	0	90
Primary Farmer	Household head's primary occupation is farming (=1)	0	1
Labor	Household head is a member of an agricultural Labor group (=1)	0	1

Cooperative	Household head is a member of an agricultural cooperative (=1)	0	1
Committee	Household head is a member of an agricultural committee (=1)	0	1
Centrality	Household head's network centrality measure	0	1
Individual Clustering Coefficient	Household head's clustering coefficient	0	1
In-Degree	Household head's total number of incoming network connections.	0	44
Out-Degree	Household head's total number of outgoing network connections	0	77
Social Preference Questions	<p>Q1 A : you receive 1000 Cf and someone else in your community receives 1000 Cf B : you receive 1000 Cf and someone else in your community receives 2000 Cf</p> <p>Q2 A : you receive 2000 Cf and someone else in your community receives 0 Cf B : you receive 1000 Cf and someone else in your community receives 1000 Cf</p> <p>Q3 A : you receive 1000 Cf and someone else in your community receives 0 Cf B : you receive 1000 Cf and someone else in your community receives 1000 Cf</p> <p>Q4 A : you receive 2000 Cf and someone else in your community receives 3000 Cf B : you receive 1000 Cf and someone else in your community receives 1000 Cf</p>		
Selfish	Household head had social preference responses Q1 = A and Q4 = A	0	1
Generous	Household head had social preference responses Q1 = B, Q2 = B, Q3 = B, Q4 = A	0	1
Inequality Averse	Household head had social preference responses Q1 = A, Q2 = B, Q3 = B, Q4 = B	0	1
Other	Household head's responses to social preference questions were not consistent with selfish, generous, or inequality averse patterns.	0	1
Dictator Game	Hypothetical amount of rice (Kg) shared with a self-selected member of the community out of a maximum of 3 Kg.	0	3

CHAPTER 5

Social Networks and Social Preferences:

A Lab-In-Field Experiment in Eastern DRC

Abstract

Considerable research points to social ties, and the social networks underlying these ties, as the underlying driver of pro-social behaviors upon which large-scale societal organization is based. However, little is known about the empirical relationship between social network position and pro-social preferences. Based on original network data and two lab-in-the-field experiments, we explore the relationship between social networks and social preferences of trust, trustworthiness and cooperation. We explore whether an individual's observed social preferences is correlated with an individual's centrality within the network structure. Our results indicate that individuals with high centrality are more trusting and more trustworthy than individuals with lower centrality. Moreover, individuals with low centrality are treated worse in these interactions—people trust them less initially, and return less money to them. Within a group context, little evidence is found of more central individuals displaying more cooperative behavior. Instead, for group cooperation, when a single monitor can observe contribution decisions, the presence of a direct link and more mutual network connections with a monitor correlates to more cooperative behavior by that individual. Our results suggest that network centrality and pro-social preferences are related but more localized network ties are more strongly correlated with pro-sociality than overall network connectedness.

Publication Status: Ross, M., P. Hofman, J. Larson, P. van der Windt, M. Voors (2017). “Social Networks and Social Preferences: A Lab-in-Field Experiment in Eastern DRC”. Working Paper.

1 Introduction

Individuals do not make economic decisions in a vacuum. Instead, individuals are embedded in rich social environments comprised of both formal and informal institutions that shape preferences and drive behavior (North, 1997). The relevance of these social institutions has been well documented: individuals both inside and outside the lab contribute to societal investments even at a direct personal cost. People assess risky investments in part based on pre-existing feelings of trust between the investor and investee (Berg, Dickhaut, & McCabe, 1995). Informal markets emerge and persist due to the interpersonal ties connecting their participants (Grief, 1993; Landa, 1981). Individuals in these situations exhibit trust and trustworthiness, the same pro-social behaviors that underpin the proper functioning of societies by generating gains from group living even absent strong formal governing institutions. While much research points to social ties, and the social networks comprised of these ties, as the underlying driver of these trust-based interactions, little is known about the empirical relationship between social network position and pro-social preferences. Based on original network data and two lab-in-the-field experiments, we explore the relationship between social networks and social preferences of trust, trustworthiness and cooperation. We explore whether an individual's observed social preferences is correlated with an individual's centrality within the network structure.

Departing from standard economic theory which holds that preferences are exogenous and stable, laboratory and lab-in-the-field experiments have shown that preferences may co-evolve with the social context (Bowles, 1998; Bowles & Polanía-Reyes, 2012; Brosig, Riechmann, & Weimann, 2007; Volk, Thöni, & Ruigrok, 2012; Voors et al., 2012). Within a social context, social preferences can vary: while there is substantial cross-cultural variation, there is also substantial within-culture variation of social preferences (Bowles, 1998; Croson & Gneezy, 2009; Fehr & Hoff, 2011; Joseph Henrich et al., 2005, 2010). One source of variance are salient social cleavages. When individuals interact with in- and out-group members, individuals tend to exhibit more pro-social behavior toward the in-group. This dynamic is shown for cleavages defined by gender (Croson & Gneezy, 2009), ethnicity and race (Benjamin, Choi, & Strickland, 2007; Fershtman & Gneezy, 2001), as well as for exogenously imposed or self-selected groupings (Ben-Ner, McCall, Stephane, & Wang, 2009; Chen & Li, 2009; Halevy, Bornstein, & Sagiv, 2008). Studies exploring

why this is the case, suggest that favorable social networks interconnecting members of an in-group are part of the answer (Habyarimana et al. 2007, Miguel & Gugerty, 2005).

Social networks may bear on pro-sociality for a number of reasons. It could be that social networks facilitate punishment, so that all who hear about misbehavior punish it. Relatedly, social networks may spread social judgment, so that misbehavior results in a worse reputation. In either case, a person's position in a social network that can transmit information or facilitate judgment could affect her incentives to behave socially or not. Whether an individual would be punished via social sanctions, or would earn a worse reputation which could affect future social capital and support, network positions from which these costs would be realized most strongly would offer a greater deterrent to anti-social behavior.

The standard approach to studying the effect of networks on behavior is to assign participants to an artificial "social network" in a lab, vary the connections in that network to vary who could punish and be punished by whom, and observe behavior. This approach is able to cleanly identify a causal effect of a network characteristic; however, the network is artificially-imposed and thus cannot illuminate whether participants draw on their own social network for these purposes. Our study differs in that we focus on the relationship between individuals' positions within their real social networks and their behavior in interactions with one another. This design sacrifices some precision in our causal effects – we cannot perfectly disentangle the mechanisms by which the network affects outcomes – in exchange for greater external validity. Specifically, we begin by measuring the social network among individuals in rural villages in the Democratic Republic of Congo. We then recruit individuals to play two lab-in-the-field experiment (an investment game and a public goods game) based on their eigenvector centrality in the village network, including some with low, middle, and high values. Individuals observe the other participants in the games, and so are free to condition their behavior on real information, including social network information. Finally, we relate individuals' eigenvector centrality in their village's social network to their observed social preferences in these games.

Eigenvector centrality is a network statistic that captures the extent to which a person is highly-connected to others in the network, and the extent to which a person's connections are themselves highly-connected to others (Bonacich, 2007; Borgatti, 2005a) Eigenvector centrality thus bears on one's access to information flowing through a network, as well as those whom one could access easily to report a bad reputation or

coordinate social sanctioning. Consequently, our design admits both post-experimental punishment mechanisms and reputation effects.

Games include a one-on-one investment game that tests trusting and trustworthy behavior, and an eight-person public goods game that tests pro-sociality in general. We add a variant to the latter which has a player serve as a monitor, observing everyone's contributions, and vary the network position of the monitor. In this way, we have information on a person's general level of trust and trustworthiness (how she behaves overall in the investment game), a person's selective level of trust and trustworthiness (how she behaves when playing certain other people in the investment game), a person's pro-social behavior in a group (how she behaves overall in the public goods game), and a person's responsive pro-social behavior in a group (how she behaves in the public goods game when someone in particular is watching).

Our results indicate that network eigenvector centrality ("centrality" for short) matters in one-on-one interactions. Individuals with high centrality are more trusting and more trustworthy than individuals with lower centrality. Moreover, individuals with low centrality are treated worse in these interactions—people trust them less initially, and return less money to them. Our results also indicate that in larger group interactions, individuals' public goods game cooperation is only weakly related to their network centrality—least central individuals cooperate slightly less. Strikingly, in these interactions, the presence of a central person who can observe all contributions has no greater effect than the presence of one of the least central villagers who can observe all contributions. In other words, central monitors are not generating fear of social sanction or reputation effects after-the-fact in these experiments, at least not to a greater extent than any other monitor. On the other hand, monitors who have a personal tie to someone in the group playing the public goods game result in *that person* contributing substantially more, and this effect holds regardless of the centrality of the monitor.

Taken together, these results suggest that network centrality is related to pro-social preferences. Individuals with high centrality scores are more trusting and more trustworthy in one-on-one investment games. At the other extreme, isolates contribute less in public goods games. However, the central are not better at improving group cooperative outcomes—their presence as monitor has no effect on the size of the public good. On the other hand, we find that direct ties or shared mutual ties improves cooperation, both for one-

on-one investments, and public good dilemmas. This suggests that the network mechanism through which social preferences are influenced direct network ties, not network position. The more pro-social behavior by centrals may be due to central players inherently having more direct ties, so social forces operating through direct ties cause them to behave more pro-socially in more game interactions. The least central have fewer direct ties, and so both behave less pro-socially and receive less favorable treatment more frequently within game play. In other words, social preferences and networks are more strongly correlated locally—with higher cooperation among pairs of peers—rather than globally—i.e. more central individuals eliciting more cooperative behavior within groups.

The existing literature on social preferences and social networks is discussed in Section 2. Section 3 offers a detailed overview of the study design. Section describes the sample, and our results are presented in Section 4. Section 5 concludes.

2 Social Networks and Social preferences

Our study focuses specifically on the relationship between the social network in which an individual is embedded and their observed social preferences.

2.1 Social preferences

While social networks capture the external relationships that define daily social interaction and access to information and resources, social preferences are the internal set of preferences determining rank order of different allocations of material benefits between oneself and others. This set of preferences includes interpersonal values of altruism, fairness, cooperation, trust, and inequality aversion. Social preferences are most often measured through laboratory experiments designed to elicit each preference under varying conditions and have become a focal point within development economics (Camerer & Fehr, 2002; Cárdenas & Carpenter, 2008). Significant research within developing country contexts have been undertaken exploring experimental designs that vary the anonymity, group composition, information, and monitoring and punishment mechanism effects on observed social preferences through lab-in-field experiments (for review see Cardenas and Carpenter, (2008)). Social preferences have been studied across cultural contexts (Cardenas, Stranlund, & Willis, 2000; Joe Henrich, 2000; Joseph Henrich et al., 2001; Jakiela, 2011) and

linked to economic outcomes of labor markets (Barr & Serneels, 2009; Carpenter & Seki, 2011; Fehr, Kirchsteiger, & Riedl, 1993), financial decisions (Karlan, 2005), markets for goods and services (List, 2005), tax systems (Alesina & Angeletos, 2003), and environmental resource management (Bouma, Soest, & Bulte, 2008) amongst others.

2.2 Social entrepreneurship and trust and trustworthiness

The theory of “social entrepreneurship” provides a mechanism by which social networks may interact with individual trust preferences. Social entrepreneurship is based on the premise that social ties that bridge structural holes are risky undertakings with potentially high social returns. Social links that bridge network sub-groups or puts an individual in the position of acting as a broker between unconnected pairs offer opportunities for economic, political, or social returns through the ability to influence or control network information channels (Burt, 1992, 1997, 2007, Lin, 1999, 2002; Sobel, 2002). These bridging opportunities thus create a potentially worthwhile pay-off for the risky and effortful endeavor of coordinating, establishing and maintaining social bridges within the network structure. In initiating engagement with a future connection and facilitating early transactions absent an interaction history, a minimum propensity to trust is required. Individuals that establish and maintain more structural holes within a network thus repeatedly engage in trust-based interactions, with successful interactions reinforcing or strengthening prior levels of trust. In smaller networks or networks with sufficient information flows, social-bridge connections may be initiated and maintained based on perceived reputations of trustworthiness held by each individual of the other. Building a reputation of trustworthiness thus potentially facilitates future bridging links as new connections more readily trust that the bridging-individual will act pro-socially (Barr, Ensminger, & Johnson, 2009).

Under the theory of social-entrepreneurship, individuals who more frequently build and maintain socially-bridging connections have higher levels of trust and trustworthiness. Empirical studies testing these theoretical predictions have used varying measures of centrality and connectedness for representing socially-bridging individuals (Barr et al., 2009) used individual scores of network betweenness within a lab-in-field investment game to measure correlations between trust and social-bridges. They found that in the contexts

of rural communities in eastern Africa and urban firms in Ghana that individuals with higher network betweenness displayed higher levels of trust.

2.3 Mechanisms: punishment and reputation for PGG

The extent to which sanctioning can facilitate higher cooperation has been studied extensively within the public goods game literature (see Fehr and Fischbacher (2004) for a review) and extended to incorporate elements of social network theory. A second additional mechanism for enforcing higher cooperation that has ties to social network theory is that of an individual's reputation as a form of social capital. Both institutions of punishment and reputation have provided insight into the complex relationships between external networks and internal preferences of cooperation using modified public good game experiments. For institutions of punishment, social sanctioning within public goods games is incorporated by allowing group members to pay a small fee to levy a larger fine on a group member for behavior that is thought to deviate from social norm. As sanctioning is at the discretion of participants, the opportunity arises for anti-social retaliation to occur which lowers the efficiency of costly punishment in facilitating greater public good provision (Boosey & Mark Isaac, 2016). When monitoring and punishment are limited to network links, network structures can be imposed exogenously on experimental groups to explore the relationship between punishment, network structures, and cooperation. Overall, more connected networks were found to outperform disconnected networks for level of contributions elicited (Carpenter, Kariv, & Schotter, 2012). The connectedness of networks was found to even be more important than the relative punishment capacity of participants for facilitating higher cooperation (Leibbrandt, Ramalingam, Sääksvuori, & Walker, 2015). Using individual level network metrics, a study by Apicelli et al. (2012) found no relationship between network in- or out-degree with cooperation. The study did find that co-operators were more likely to be connected to other co-operators and vice-versa for defectors. This separate clustering of co-operators and defectors is thought to demonstrate social sanctioning of defectors, allowing only co-operators to realize gains from public goods and thus facilitating further cooperation (Apicella et al., 2012).

An existing theory behind the use of sanctioning for facilitating cooperation is that sanctioning requires defectors to be 'findable'. Higher social connectedness facilitates sanctioning of defectors, increasing the probability of punishment in retaliation for anti-social behavior (Miguel & Gugerty, 2005). This 'findability'

effect is found to hold in revealed games even when punishment is not possible within the experiment. Since in-group members are likely to be more closely socially connected outside of experiments, participants are more cooperative due to the risk of post-experiment sanctioning of defection within the experiment (Habyarimana, Humphreys, Posner, & Weinstein, 2007; Haley & Fessler, 2005).

The second mechanism for increasing cooperative behavior within public good experiments is the use of reputation as a form of social capital. An individual's reputation can be conceptualized as a tradeable commodity, with positive reputations increasing the probability of receiving social assistance when required, even from strangers, as long as the reputation is known (Milinski, 2016). Good reputations are built by making pro-social investments in social group members while anti-social behavior results in reputational decline. Thus costly investments in social group members can have future returns in the form of social assistance, providing an informal insurance system. Experimental evidence testing the effectiveness of reputation effects in strengthening cooperation within public good provision found that the presence of an artefactual reputation mechanism resulted in sufficiently high contributions for all group members to have higher overall returns relative to rounds played without a reputation mechanism (Milinski, Semmann, & Krambeck, 2002).

Similar to punishment mechanisms, the effectiveness of reputational mechanisms is strongly linked to underlying social group identities and social network parameters. Any individual's reputation is constructed as the cumulative of pro- or anti-social behaviors observed and spread through informational flows within their respective social network. Therefore, the reputation of an individual is closely tied to social gossip, allowing for perceived reputations of others to be known by strangers based on reputation information transmitted by gossip through social networks (Sommerfeld, Krambeck, & Milinski, 2008). A theoretical study by Larson, set within the context of inter- and intra-ethnic cooperation, explored how information flows affect cooperation. The theoretical predictions were that wider and more rapid information flow paths facilitated more cooperative behavior (Larson, n.d.). Thus the more reachable individuals are, the more likely their social network structure will foster higher overall cooperation by individual network members. While the study framed word-of-mouth communication of poor behavior as a form of peer sanctioning, it captures the same traits of a negative reputation effect. Evidence suggests that members of social networks can, to

an extent, identify individuals that are better situated for gossip channels (Banerjee, Chandrasekhar, Duflo, & Jackson, 2013).

When reputation mechanisms are contrasted with punishment mechanisms within artefactual experiments, reputation mechanisms were found to be both more effective and more efficient relative to punishment (Wu, Balliet, & Lange, 2016). This finding was driven by the non-costly reputation effects increasing both individual contribution levels and overall earnings at sustained levels throughout all rounds played. Comparatively, punishment mechanisms increased individual contribution levels only in later rounds and the costliness of punishment resulted in overall lower collective earnings. However when both reputation effects and punishment were both available within game play, costly punishment persisted only at low levels and was concentrated against free riders, resulting in an interaction effect that boosted overall efficiency of player earnings (Rockenbach & Milinski, 2006). Reputation effects are thus found to be a powerful and low-cost alternative to punishment. Within our study both reputation and punishment effects may indirectly effect player decision making through revealed game play, however neither mechanism is directly incorporated within the study design. Providing information on both the identity of the contributor and the amount contributed results in substantial increases in giving to the public good (Andreoni & Petrie, 2004), however reputation effects cannot be disentangled from ex-post social sanctioning outside the artefactual experiment.

2.4 Contribution to the Literature

We contribute to this literature in three ways.

First, we use network data at the community level. By utilizing community social network data we are able to explore how an individual's position within their most relevant social network is related to observed social preferences. By using real-world networks we contextualize decisions to capture both within game and ex-post social forces that potentially relate to social preference behavior. While this compromises the identification of causal relationships, it provides an indication of how social preferences and social networks interact within real-world networks.

Second, use of full network data allows us to calculate the centrality of individuals within their communities rather than just within experimental participant groups. Numerous measures of centrality for capturing the importance of a network member for the flow of information and resources exist. Each has its own functional forms that uniquely captures the varying traits or characteristics inherent within flow patterns of different resources or information (Borgatti, 2005a). Eigenvector centrality is often considered a measure of an individual's influence within a network. An individual is considered more influential the better connected they are to central individuals within the network, so eigenvector centrality is a recursive measure of connectedness (Bonacich, 2007). This recursive function is solved by using degree centrality weightings for an iterative estimation approach. While eigenvector centrality is popularly considered a measure of influence, it can also be thought of as the probability of an individual participating in any resource flows that are able to travel via unrestricted walks. In other words, for resources that can follow any number (infinite divisibility) of any links (unrestricted) between network nodes, eigenvector centrality provides an approximate likelihood of a given node receiving and passing along the flowing resource. Information is the most common resource that flows through networks both divisibly and unrestricted, making eigenvector centrality an appropriate measure for capturing the extent to which an individual participates, influences, and potentially controls, information flows within their network (Borgatti, 2005b).⁵⁸ Our study provides insights on whether more centrally positioned individuals are, on average, more pro-social in their preferences relative to less central individuals. Such relationship would indicate that social preferences and social networks are co-evolutionary in their formation, although this study is not designed to explore potential causal or directional relationships.

Third, by combining network data with revealed game play we are able to explore the responsiveness of social preferences to localized social forces of network links and shared network connections. This allows us to gain a better understanding of whether individuals' social behavior is more responsive to the threat of wider information flows (e.g. a greater number of more socially-distant individuals) versus localized forces (e.g. directly connected individuals) in facilitating pro-sociality.

⁵⁸ Banerjee et al. (2013) devise an information-diffusion specific measure of centrality that they term diffusion centrality. This measure is closely related in mathematical form to eigenvector centrality and so we persist with eigenvector centrality as the determining network parameter of interest within our study design.

3 Experimental Design

To test our claims we combine social network data of household heads with two lab-in-field experiments. To collect data for this study we visited each village two times. The first visit entailed a household survey with all household heads to collect social network information as well as socioeconomic data. The second visit occurred approximately one month later, during this visit we implemented our lab-in-the-field experiments.

3.1 *Obtaining network information*

Our research assistants first conducted a full village census of all heads of household in the village, during which the head of household's full name, age, and gender were recorded along with the names of any other adults in the household. Upon completion of the census, each household head was individually interviewed.

The census consisted out of two parts. The first part collected information about basic socio-economic characteristics, such as demography as well as income and related questions. In the second part we collected social network data. Specifically, we aimed to obtain data on three types of network. The first is *family network*. Whether the head of the household is biological family with any member within the other household.⁵⁹ Second, we ask about *field-neighbor network*. Whether the head of the household's fields borders a field owned by any member of the other household. Third, we measure the *agriculture network*. Whether the head of the household discusses agricultural related topics with anybody else in the other household. These three networks were carefully selected. Pilot studies which included a larger number and greater variety of relationship questions indicated that these three relationships captured the most network variation while minimizing survey fatigue. We focus on these three networks because they are most closely associated to our interest. That is, Kendzior, Zibika, & Voors (2015) found that these three networks are the three predominant channels via which agricultural resources were shared. We conducted a number of pilots that

⁵⁹ Specifically, we use whether the other person is biologically related to a maximum of the third degree (this is a well-understood term in Congo). This does not include the wife's family; it has to be through descent.

included more networks.⁶⁰ Pilot results showed that these three dimensions were the most distinct from one another and thus captured maximum variation while minimizing the number of network survey questions.

Specifically, these relationships were elicited by first constructing a complete roster for the entire community of all household heads, wives of household heads, and the presence of other adult members (greater than 16 years of age) within the household. With this roster, each relationship type was its own survey question in which the interviewer informed the respondent that each name from the roster would be read aloud. For each name read aloud the respondent would indicate ‘yes’ or ‘no’ on whether there existed between them the relationship in question.⁶¹ Household heads and their wives had their names read off the roster individually, however all other household adult members were grouped into a single category and phrased “Are you [*connection type*] with any other adults within this household?”

Once network data along the three connections was collected for all household heads the data was first collapsed to the household level. If any connection exists between a household head and a member of another household, the two households are said to be linked. The result was three separate network graphs with each node representing a single household and each connection being unweighted and undirected. Secondly, we collapsed these three to a single network graph representing all three connection types. These three network graphs were reduced to a single composite network graph in which the presence of a link between two households along any connection type is treated as an undirected and unweighted link within the composite graph. The result is a single network graph for each community in which each node is a household and each link represents the presence of any measured connection being declared from one or both of the household nodes. The eigenvector centrality score for each household was calculated using the respective community’s composite social network graph. To increase comparability of a household’s eigenvector centrality score across communities, eigenvector centrality scores are normalized by community size measured as the total number of resident households.

⁶⁰ In total we conducted three pilots. The other networks were friends (the problem was that everyone was everyone’s friend), and work on another person’s farm (that overlapped with the other networks).

⁶¹ This approach was considerably more time intensive than alternative strategies in which names were offered up by the respondent for each relationship but overcame potential measurement error arising from memory-based recall. This approach also offered the benefit of reducing misidentification of individuals in a context with frequently repeated names.

The participants in our lab-in-the-field experiments were selected based on the rank order of households' eigenvector centrality scores. We selected the six household heads with the highest centrality score, whom we call "Centrals". We also selected those with the lowest centrality score, the "Isolates". Finally, we also selected the six heads of households with the median centrality scores, the "Middlings".

About one month later after the first visit, we revisited each village to conduct two lab-in-the-field experiments with selected participants: the investment game and the public goods game. We discuss these now.

3.2 *Measuring trust and trustworthiness*

To measure trust and trustworthiness we make use of the fully-revealed investment game, following Berg et al. (1995). In this game, two participants are partnered together as a Player 1 and a Player 2. From an initial endowment of 15 tokens Player 1 decides how many, if any, they wish to share with their partner. Each token was valued at 100 Congolese Franc. The total endowment (1,500 FC) is equivalent to a day's work. Any token amount that is shared is tripled and given to Player 2. Player 2 is then given the opportunity to return any amount of this received money back to Player 1. Both Player 1 and Player 2 are informed of the identity of their partner. The number of tokens shared by Player 1 captures the level of *trust* Player 1 has in Player 2. The number of tokens returned by Player 2 is a measure of Player 2's *trustworthiness*. We implement this experiment following a round-robin game, where each participant is paired one time with the other participants. In other words, each participant is 17 times Player 1, and 17 times Player 2.⁶²

3.3 *Measuring cooperation*

After the investment game, participants took part in a public goods game. Specifically, the 18 participants were divided into three groups of six members each. Each group consisted of two Centrals, two Middlings, and two Isolates. Each participant received 15 tokens and had to decide what number of tokens to contribute to a public account. Once all contribution decisions were made, the public account was doubled and divided equally amongst all group members regardless of contribution amount. While the group composition is fully revealed for all variations of the public goods game played, the privacy of the individual

⁶² The order of partners was randomized in advance by ID number. As ID numbers were not randomized within the tertiles, there are some potential ordering effects present.

contribution decisions varied from fully private, public to a single identified group member, to fully public. Each participant plays four rounds of the game. This setup is the workhorse to measure an individual's incentivized cooperation subject to perceived group expectations of social norms (Andreoni & Petrie, 2004).

3.4 Cooperation under monitoring

To learn about the importance of monitoring and an individuals' position within the network on contribution behavior, each participant played five variations of the above-described game. *Variation 1* is the basic setup as described above with all information remaining private. *Variation 2* is a fully revealed public goods game. In contrast to the previous variation, all individuals' contribution decisions are made public. *Variation 3* is similar to variation 1 but a monitor is selected. The identity of the monitor is made public to the group prior to play commencing. The individuals' decisions are made known only to the monitor. In this variation, the monitor is a central participant. *Variation 4* is the same as the previous variation, but the monitor is a middling. *Variation 5* is the same as the previous variation, but the monitor is a randomly selected isolate. Our design is illustrated in **Table 1**Error! Reference source not found..

Table 1 Public Goods Game Variations

Variation	Label	Monitor Information	Public Information
1	Fully anonymous	NA	Public account
2	Fully revealed	NA	Player contributions & Public account
3	Central monitor	Player contributions	Public account
4	Average monitor	Player contributions	Public account
5	Isolate monitor	Player contributions	Public account

The order in which participant played the five variations was randomized to avoid ordering effects. In addition, participants were randomly assigned to each of the three groups. As part of variations 3 to 5, each corresponding participant fulfils the role of monitor for two rounds of play. In other words, in variation 3, one central participant will be the monitor for two rounds. After two rounds, the other central participant

will be monitoring for the remaining two rounds. The order of monitors in each variation was not randomly assigned.⁶³

At the conclusion of both the investment and public goods games, each participant received compensation based on their decisions made with one randomly selected partner-pairing from the investment game (paid for rounds played with that partner as both Player 1 and Player 2) plus the payouts from one randomly selected rounds of public goods game play across all five variations and all retained silence tokens. Participants were paid out for both games simultaneously at the end of the day so participants were not aware of their realized earnings at any point during game play.⁶⁴

3.5 Estimation strategy

To learn about social network related determinants of trust and trustworthiness, we estimate the following equation:

$$Y_{ijz} = \beta_1 \text{Centrality}_{iz} + \beta_2 \text{Centrality}_{jz} + \beta_3 \mathbf{X}_{ijz} + \gamma_r + \alpha_z + \varepsilon_i \quad (1)$$

where Y_{ijz} is the number of tokens contributed by individual i with partner j in community z . Thus for trust, i represents Player 1 and j represents Player 2, while for trustworthiness the opposite is true. Centrality_{iz} is an indicator variable for the centrality type of individual i in community z , where $\text{Centrality}_{iz} \in \{\text{central}, \text{middling}, \text{isolate}\}$. Centrality_{jz} is the same for their partner. \mathbf{X}_{ijz} is a vector of individual characteristics for both Player 1 and Player 2. Included in this vector is gender, migrant status, age, literacy, and highest education level attained (none, primary, secondary, tertiary). Finally, γ_r are round fixed effects, α_z are village fixed effects in order to control for localized social norms that may affect general expectations of social behavior, and ε_i is the residual error term clustered at the individual decision-maker level (Player 1

⁶³ Specifically, the order of monitoring determined by ID number from lowest to highest within each group. This opens up potential anchoring effects, in that within variation, the more central of the two monitors will play first. We do not expect this to be a big problem as the monitors within variation have a nearly indistinguishable centrality. If there are anchoring effects and more central individuals are better monitors then the coefficient should provide an upper bound of the effect.

⁶⁴ In addition, to minimize sharing of private information between participants, each player was incentivized not to share information regarding their decisions throughout the day with ‘silence tokens’. Silence tokens represented additional bonuses of 50 Congolese francs to be earned by not discussing private information and were earned for each round of play. If players were overheard sharing information that was meant to remain anonymous or private, silence tokens for that round were confiscated and noted on record sheets.

for outcomes of trust and Player 2 for outcomes of trustworthiness). Note that in the regressions related to trustworthiness, we also control for the number of tokens received by Player 2 from Player 1.

Previous studies (Habyarimana et al., 2007; Miguel & Gugerty, 2005) have highlighted the role of ‘findability’ for more effective social sanctioning, thus providing an incentive for more closely connected individuals to behave pro-socially. As a result, we rerun equation (1) but where we replace the participants’ centrality measures ($Centrality_{iz}$ and $Centrality_{jz}$) with relational network characteristics between the two partners. Specifically, we measure the relationship between two partners in three ways. First, the difference in eigenvector centrality score between individual i and partner j . Second, an indicator for the fewest number of links connecting i and j . Third, the proportion of direct network connections shared by i and j within their overall community social network.⁶⁵

To learn about the role of social network related factors for contributions in the public goods game, we estimate the following equation:

$$Y_{igz} = \beta_1 Centrality_{igz} + \beta_2 Group_{gz} + \beta_3 Var_{igz} + \beta_4 \mathbf{X}_{iz} + \beta_5 \mathbf{M}_{igz} + \gamma_r + \alpha_z + \varepsilon_i \quad (2)$$

where Y_{igz} measures the proportion of the initial endowment contributed to the public account by individual i of experimental group g in community z . $Centrality_{igz}$ is again the centrality type of individual i . We also include $Group_{gz}$, which captures the proportion of possible direct links between public goods game group g members observed within community z as a measure of localized group density. Var_{gz} indicates the variation, which captures the extent to which information is publicized. \mathbf{X}_{iz} includes the same set of individual characteristics as in equations 1 and 2.⁶⁶ \mathbf{M}_{igz} is an indicator that individual i is the designated monitor. Finally, γ_r are round fixed effects, α_z are village fixed effects, and ε_i is the residual error term clustered at the individual level.

⁶⁵ Due to high correlation between direct link and share of network connections, we split these regressions out separately, controlling for difference in eigenvector centrality within each. This circumvents concerns of collinearity within our estimation results.

⁶⁶ While the Investment Game was conducted prior to the Public Goods Game in all communities visited. However, participants were not informed of their earnings until after all game play was concluded. Including a measure of average Investment Game earnings for individual i in the Public Goods Game analysis does not affect estimated results.

Similar to the investment game we also explore the relationship between the participant and the monitor. As a result, we rerun equation (2) but add two measures related to the relationship between the participant and the monitor. First, we use a measure captures the proportion of shared network connections within the larger community \mathcal{N} between individual i and monitor m of group g . Second, we use an indicator variable taking value 1 if individual i and monitor m are directly connected within the community social network.⁶⁷

4 Data and sample

Before moving to the results in the next section, we first introduce our data and the participants.

4.1 Data and attrition

For the census upon which we base our network measures we took every effort was taken to ensure a high level of response in our census, which translates into a very high response rate of 97%.⁶⁸ In total, we aimed to collect trust and trustworthiness data from 720 participants (18*40) and 12,240 pairings (18*17*40). In total, this study builds on data from 11,774 investment game pairs, played by 706 participants. For the public goods game we aimed to collect for 720 participants over 14,400 rounds (4*5*720). In total, we build on data from 712 participants, and 14,192 rounds. Lost data was due to some participants game ID's being incorrectly entered during game play, resulting in lab-in-field data not able to be linked to survey data.

4.2 Manipulation check: Does network position actually mean something?

We are interested in differences due to different positions of participants in the social network. We now check if being a central, a middling or isolate, actually correspond to other characteristics. **Table 2** present the differences. In exploring the variation in socio-economic variables across centrality types, we conduct a difference in means test between each of the three centrality groups. We find several large differences, and some similarities. Centrals are more likely to be literate than Middlings, and Middlings are more literate than Isolates. This monotonous sloping trend is consistent for all differences we find. Lower centrality is associated with being a migrant, and being a female-headed household. We also see that more central

⁶⁷ If individual i = monitor m than *Direct Link* = 0.

⁶⁸ When a head of household was absent, we returned to the village for a second time (most often that weekend). If the head of household was still absent, a replacement within the household was asked to stand-in for the household head. Representatives were asked to respond to all question from the perspective of the household head.

individuals have a higher education level, though this only holds significantly for primary and secondary education. This is probably because for higher levels we have very little variation (only very rarely had people followed more than secondary education).

Table 1 Description Game Participants

VAR	Overall	Central	Middling	Isolate	Difference in Means Test (p-value)		
					<i>Central Middling</i>	<i>Central Isolate</i>	<i>Middling Isolate</i>
Age (years)	47.632 (0.961)	47.867 (1.198)	48.742 (1.250)	46.274 (1.532)	0.540	0.381	0.125
Literacy (=1)	0.424 (0.034)	0.515 (0.042)	0.424 (0.041)	0.332 (0.044)	0.038	0.000	0.033
Migrant (=1)	0.311 (0.017)	0.129 (0.023)	0.305 (0.034)	0.504 (0.035)	0.000	0.000	0.001
Female (=1)	0.324 (0.022)	0.167 (0.025)	0.326 (0.037)	0.483 (0.045)	0.001	0.000	0.008
<i>Highest Edu Attained</i>							
<i>Primary</i>	0.307 (0.023)	0.358 (0.036)	0.275 (0.025)	0.286 (0.033)	0.020	0.099	0.755
<i>Secondary</i>	0.213 (0.027)	0.262 (0.037)	0.220 (0.035)	0.154 (0.031)	0.279	0.006	0.063
<i>Tertiary</i>	0.014 (0.006)	0.017 (0.008)	0.013 (0.007)	0.013 (0.007)	0.679	0.688	0.986
Inequality Aversion	0.615 (0.028)	0.592 (0.047)	0.581 (0.042)	0.675 (0.036)	0.832	0.163	0.064
N	710	240	236	234			

5 Results

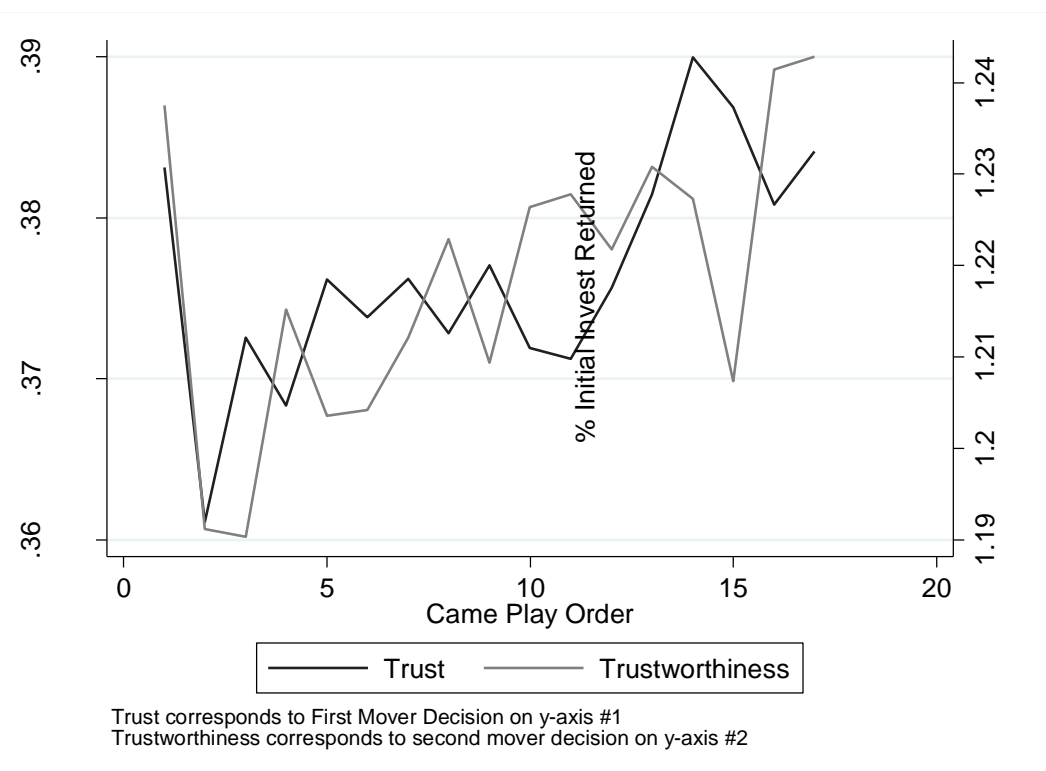
In this section we present the results related to trust and trustworthiness. We then present our results related to contributions in the public goods game.

5.1 The social network and trust and trustworthiness

Trust is measured as the proportion of the initial endowment that is shared by Player 1 to Player 2. We find that as Player 1s, Centrals share, on average, 41% of their initial endowments with Player 2 while Middlings and Isolates share, on average, 38% of their initial endowments with Player 2. Trustworthiness is measured as the percentage of the initial investment returned by Player 2. We find that Centrals returned an average of 129%, compared to 119% from Middlings and 121% from Isolates.

Each participant played the investment game 34 times: 17 times as Player 1, and 17 times as Player 2. The round in which participants played can have an influence on game play. **Figure 1** shows that moving from the first round of play to the second round there is a decrease in both trust (3%) and trustworthiness (5%). This drop is recovered, however, over the remaining 15 rounds of play for both Player 1 decisions of trust and Player 2 decisions of trustworthiness. We randomly assigned the order so as not to be worried about ordering effects.

Figure 1 Trust And Trustworthiness By Order Of Play



Our results are reported in **Table 2**. Column (1) in **Table 2** presents results from equation (1). We find that central individuals are on average sending four percentage points more of the initial endowment than middling participants within the investment game. In testing the statistical difference in coefficient estimates between central and isolate trust behaviors, the difference is found to be statistically significant at the 10%. However, there is no significant difference in trust behavior between Isolates and Middling individuals. This result indicates that while centrality is correlated to the social preference of trust, the correlation is not linear across all centrality types. Instead, the positive correlation occurs only among the most centrally positioned individuals within a social network.

Table 3 Regressions Results Related To Trust

VARIABLES	(1) Trust	(2) Trust	(3) Trust	(4) Trustworth.	(5) Trustworth.	(6) Trustworth.
<i>Player 1</i>						
Central	0.036** (0.014)			-0.012 (0.012)		
Isolate	0.009 (0.015)			-0.025** (0.011)		
<i>Player 2</i>						
Central	0.001 (0.003)			0.067* (0.034)		
Isolate	0.002 (0.003)			-0.016 (0.034)		
Difference Centrality (P1 – P2)		-0.026** (0.010)	-0.026** (0.010)		0.056** (0.024)	0.056** (0.024)
Direct Link		0.020*** (0.006)			0.044*** (0.015)	
Shared Connections			0.120** (0.052)			0.200 (0.124)
P2 Endowment				0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Constant	0.405*** (0.042)	0.416*** (0.039)	0.409*** (0.040)	1.360*** (0.094)	1.340*** (0.093)	1.332*** (0.094)
Observations	11,614	11,614	11,614	11,517	11,517	11,517
R-squared	0.204	0.204	0.204	0.117	0.116	0.116
Socio-Econ Indicators	P1 & P2	P1 & P2	P1 & P2	P1 & P2	P1 & P2	P1 & P2
Village Fixed Effects	Y	Y	Y	Y	Y	Y
Round Fixed Effects	Y	Y	Y	Y	Y	Y
<i>p-value for F-test</i>	0.0948	<i>p-value for F-test</i>		0.642		
P1 Central = P1 Isolate		P2 Central = P2 Isolate				

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered by Player 1 for trust regressions and Player 2 for trustworthiness regressions are reported in parentheses. Trust and Trustworthiness (Trustworth.) measured as proportion of initial endowment and initial investment, respectively

Next we explore outcomes on trust based on relational network parameters between first and Player 2s. Column (3) and (4) of **Table 3** show that relationship variables are important. As the difference in eigenvector centrality scores between Player 1 and Player 2 increases, trust behavior decreases by over two percentage points. This translates to a reduction in mean trust behavior of 6%. As this difference in centrality measure is directional, a greater difference in centrality indicates that Player 2 is more central than Player 1. This result suggests that more isolated individuals are less willing to trust in more central individuals within their community. Network relations that increase trust behavior include the first and Player 2 being directly connected (Column (3)) and sharing a greater number of connections within the community social network (Column (4)). Interestingly, adding a single additional shared link is correlated more strongly with resulting

trust behavior (12%) than having a direct network tie (2%) suggesting that the reputation effect is more influential than the findability effect⁶⁹.

Next, we move to trustworthiness. Results are reported in Columns (4) – (6) of **Table 3**. Central individuals are more likely to return a higher proportion of the initial investment by 6.7 percentage points (5%) to Player 1s compared to Middling individuals, although this difference does not hold statistically between Centrals and Isolates. When Player 1 is an Isolate, individuals are less generous in their return on investments, returning on average 2.5 percentage points less (-2%), however this decreased trustworthiness is not observed for Middlings or Centrals.

We find that the relationship between Player 1 and Player 2 is again important. Results are presented in Column (5) and Column (6). Both a direct connection between Player 1 and Player 2 and a greater difference in centrality are found to affect trustworthiness behavior (Column (5)). More central Player 2s are more generous in their returns on investment to more isolated Player 1s, as seen in the positive and significant coefficient estimate on difference in eigenvector centrality. When a Player 2 is directly connected to their Player 1 partner, they will return nearly 4 percentage points (3%) more compared to pairings in which no network link is present. No relationship between more mutual connections and trustworthiness is found to be statistically significant (Column (6)). This is in contrast to trust, where the shared network connections were found important for higher trust levels.

5.2 *Public Goods Game Results*

The average contribution to in the public goods game is 45% with little variation by centrality. Results for equation (2) regression analysis are reported in **Table 4**. Columns (1) shows that there are few difference across treatment variations and player centrality scores. Isolates contribute less (4%) on average compared to Middlings, significant at the 10% level. We find no difference in average contributions made by Centrals compared to Isolate. Furthermore, we observe no effect of monitor centrality on average contributions, indicating that connectedness of a monitor is not strong enough facilitate cooperation by triggering potential reputation or findability effects. There are no differences in average contribution levels based on differences

⁶⁹ As our mean village size is 64 households, a single additional shared link corresponds on average to approximately a 1% increase in proportion of shared links.

in group density, suggesting that greater connection density within the group is also not a strong enough mechanism for increasing cooperative behavior.

Table 4 Regressions Results Related To Trustworthiness

VARIABLES	(1) Cooperation	(2) Cooperation	(3) Cooperation
<i>Centrality Type</i>			
Central	-0.005 (0.012)		
Isolate	-0.018* (0.011)		
<i>Variation</i>			
2 – Fully Revealed	0.000 (0.008)		
3 – Central Monitor	-0.002 (0.008)	0.006 (0.011)	0.001 (0.011)
4 – Middling Monitor	-0.007 (0.008)		
5 – Isolate Monitor	0.000 (0.008)	0.007 (0.008)	0.008 (0.008)
<i>Relation to Monitor</i>			
Difference in Centrality (Monitor – Contributor)		-0.008 (0.016)	-0.004 (0.015)
Shared Connections		0.087** (0.041)	
Direct Link			0.022*** (0.008)
Group Density	-0.016 (0.050)	-0.005 (0.064)	0.008 (0.064)
Constant	0.483*** (0.041)	0.450*** (0.042)	0.437*** (0.043)
Observations	12,802	7,708	7,708
R-squared	0.270	0.288	0.288
Variations	ALL	3 - 5	
Socio-Econ Indicators	Y	Y	
Village Fixed Effects	Y	Y	
Round Fixed Effects	Y	Y	
<i>p-value</i> of F-test: Central = Isolate	0.291		
<i>p-value</i> of F-test : Variation 4 = Variation 5		0.400	0.340

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Robust standard errors clustered at the individual level are reported in parentheses. Column (2) and (3) only include non-monitored variations as no difference in centrality, direct link, or shared connection relationships exist.

Similar to our investment game analysis, we also explore results based on the relationship between participants. Columns (2) and (3) of **Table 4** reports results for the relationship between the participant and the monitor. Similar to results observed within the investment game, relational network parameters between individuals and the monitor have a stronger relationship with cooperative behavior than individual network parameters of either the group member or the monitor individually. From column (2) we find that when a

group member shares a larger proportion of community network connections, they increase their contribution levels by an average of nearly 9 percentage points, the equivalent of a 20% increase in average contributions. Additionally, a direct link with the monitor increase average contribution levels by 2.2 percentage points. This result reinforces the result observed in the investment game that the reputation effect may have a stronger effect than the findability effect – indicating that individuals are more responsive to potential information spread of their behavior to their localized social connections than they are responsive to potential direct sanctioning by a single individual to whom they are directly linked.

6 Conclusion

Building on literature exploring origins and drivers of social preferences, we explore relationships between social network positioning and social behaviors in two lab-in-field games. We focus on the network eigenvector centrality of individuals within investment game and public goods game. We employ a round robin styled investment game with fully revealed play to explore how behaviors of trust and trustworthiness are tied to first and Player 2 centrality measures. In exploring cooperation, we employ a public goods game varying the presence of a monitor, the centrality of the monitor and compare to non-monitored revealed and anonymous public goods game designs. These different variations allow us to explore relationships between centrality scores and cooperation as well as the potential reputational or ex-post punishment effects that may vary in magnitude by the network centrality of monitors. Using household network data collected from 40 communities game participants are selected as the most central, least central, and those in the middle, based on calculated eigenvector centrality score.

Our investment game results indicate that centrality is positively but not uniformly correlated with social preferences of trust, trustworthiness and cooperation. While Central individuals displayed higher levels of trust and trustworthiness behaviors in the investment game, Isolate individuals displayed less cooperative behavior in the public goods game relative to the other two centrality types. These results indicate that there are shaping forces tying an individual's social preferences and positioning within their community social networks although our results are unable to isolate the direction of that relationship. Determining whether more pro-social individuals become more central or whether more central individuals develop more pro-social preferences is beyond the scope of this paper. However it is clear that the two are inextricably linked,

providing evidence supporting existing theoretical literature on relationships between pro-social behaviors and centrality within social networks.

Expanding beyond traits of an individual to exploring mutual network traits between paired individuals in social transactions highlights the more substantial role of relational network indicators. In the bilateral investment game experiment, partnered individuals exhibited higher trust and trustworthiness levels when they shared a greater proportion of mutual connections within their social network and had a direct relationship tie. Similarly, in the multilateral public goods game environment, when a single monitor was informed of the other group member contribution levels, the group members responded to the monitor's presence with greater cooperation when they were directly connected via a network link with the monitor and had a greater proportion of mutual connections with the monitor within the larger community network.

These latter two findings suggest that pro-social enforcement mechanism of ex-post experimental findability and reputation effects are strong enough to elicit greater pro-social behaviors but only at a very localized level. An individual being more central within a social network is not sufficiently strong enough of a mechanism to facilitate pro-social behavior. Instead these reputation effects and findability effects are occurring only when the informed party (either the partner in the investment game or the monitor in the public goods game) has a more direct influence over the local connections an individual holds within their community network. This direct influence may be either through mutual links or a direct link with the individual themselves. Our study is unable to determine whether this localization of reputation and findability effects is due to individuals being unaware of the relative centrality levels of individuals to whom they are only distantly connected, or share few mutual connections with, or whether it is due to the findability and reputation mechanisms by more distant individuals being of less concern when making social versus individualistic allocation decisions. Research by Banerjee et al. (2013) on the capabilities of community members to approximate centrality levels of peers within their social networks, even when distantly connected. Furthermore, our study is conducted in small communities of under 100 households, making it improbable that knowledge constraints are a major factor in play within our study context. With these two considerations in mind, we hypothesize that the reduced pro-social behavior is not attributed to

insufficient network knowledge but instead weaker social linkages resulting in reduced effectiveness of findability and reputation effects in eliciting more pro-social behavior.

References

- Alesina, A., & Angeletos, G.-M. (2003). Fairness and Redistribution: U.S. versus Europe. *NBER Working Paper Series*, 9502(2).
- Andreoni, J., & Petrie, R. (2004). Public goods experiments without confidentiality: A glimpse into fund-raising. *Journal of Public Economics*, 88(7–8), 1605–1623. [https://doi.org/10.1016/S0047-2727\(03\)00040-9](https://doi.org/10.1016/S0047-2727(03)00040-9)
- Apicella, C. L., Marlowe, F. W., Fowler, J. H., & Christakis, N. a. (2012). Social networks and cooperation in hunter-gatherers. *Nature*, 481(7382), 497–501. <https://doi.org/10.1038/nature10736>
- Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013). The diffusion of microfinance. *Science*, 341(6144), 1236498.
- Barr, A., Ensminger, J., & Johnson, J. C. (2009). Social networks and trust in cross-cultural economic experiments. In K. S. Cook (Ed.), *Whom Can We Trust? How Groups, Networks and Institutions Make Trust Possible* (pp. 65–90). Russell Sage Foundation.
- Barr, A., & Serneels, P. (2009). Reciprocity in the workplace. *Experimental Economics*, 12(1), 99–112. <https://doi.org/10.1007/s10683-008-9202-8>
- Ben-Ner, A., McCall, B. P., Stephane, M., & Wang, H. (2009). Identity and in-group/out-group differentiation in work and giving behaviors: Experimental evidence. *Journal of Economic Behavior and Organization*, 72(1), 153–170. <https://doi.org/10.1016/j.jebo.2009.05.007>
- Benjamin, D. J., Choi, J. J., & Strickland, A. J. (2007). *Social Identity and Preferences* (NBER Working Paper Series).
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity and Social History. *Games and Economic Behavior*, 10, 122–142.
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29, 555–564. <https://doi.org/10.1016/j.socnet.2007.04.002>
- Boosey, L., & Mark Isaac, R. (2016). Asymmetric network monitoring and punishment in public goods experiments. *Journal of Economic Behavior and Organization*, 132, 26–41. <https://doi.org/10.1016/j.jebo.2016.09.015>
- Borgatti, S. P. (2005a). Centrality and network flow. *Social Networks*, 27, 55–71. <https://doi.org/10.1016/j.socnet.2004.11.008>
- Borgatti, S. P. (2005b). Centrality and network flow. *Social Networks*, 27(1), 55–71. <https://doi.org/10.1016/j.socnet.2004.11.008>
- Bouma, J., Soest, D. Van, & Bulte, E. (2008). Trust, Trustworthiness and Cooperation : Social Capital and Community Resource Management. *Journal of Environmental Economics and Management*, 56(2), 1–30.
- Bowles, S. (1998). Endogenous preferences: The cultural consequences of markets and other economic institutions. *Journal of Economic Literature*, 36(1), 75–111. <https://doi.org/10.2307/2564952>
- Bowles, S., & Polanía-Reyes, S. (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature*, 50(2), 368–425. <https://doi.org/10.1257/jel.50.2.368>
- Brosig, J., Riechmann, T., & Weimann, J. (2007). *Selfish in the end?: An investigation of consistency and stability of individual behavior*. MPRA.
- Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Harvard University Press.
- Burt, R. S. (1997). Contingent Value of Social Capital. *Administrative Science Quarterly*. <https://doi.org/10.1037/0011646>
- Burt, R. S. (2007). *Brokerage and Closure: An Introduction to Social Capital*. Clarendon lectures in management studies series. Oxford University Press. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Camerer, C. F., & Fehr, E. (2002). Measuring Social Norms and Preferences using Experimental Games : A Guide for Social Scientists Measuring social norms and preferences using experimental games : A guide for social scientists. *Research in Economics*, (97), 55–95. <https://doi.org/10.1093/0199262055.003.0003>
- Cárdenas, J. C., & Carpenter, J. P. (2008). Behavioural Development Economics: Lessons from Field Labs in the Developing World. *The Journal of Development Studies*, 44(3), 311–338. <https://doi.org/10.1080/00220380701848327>
- Cardenas, J. C., Stranlund, J., & Willis, C. (2000). Local environmental control and institutional crowding-out. *World Development*, 28(10), 1719–1733. [https://doi.org/10.1016/S0305-750X\(00\)00055-3](https://doi.org/10.1016/S0305-750X(00)00055-3)
- Carpenter, J., Kariv, S., & Schotter, A. (2012). Network architecture, cooperation and punishment in public good experiments. *Review of Economic Design*, 16(2–3), 93–118. <https://doi.org/10.1007/s10058-012-0120-z>
- Carpenter, J., & Seki, E. (2011). Do social preferences increase productivity? Field experimental evidence from fishermen in Toyama Bay. *Economic Inquiry*, 49(2), 612–630. <https://doi.org/10.1111/j.1465-7295.2009.00268.x>
- Chen, B. Y., & Li, S. X. (2009). Group Identity and Social Preferences. *The American Economic Review*, 99(1), 431–457.
- Crosen, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, 47(2), 448–474. <https://doi.org/10.1257/jel.47.2.448>
- Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, 8(4), 185–190. <https://doi.org/10.1016/j.tics.2004.02.007>
- Fehr, E., & Hoff, K. (2011). INTRODUCTION : TASTES , CASTES AND CULTURE : THE INFLUENCE OF SOCIETY ON PREFERENCES *. *The Economic Journal*, 121(November), 396–412. <https://doi.org/10.1111/j.1468-0297.2011.02478.x>

- Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does Fairness Prevent Market Clearing ? An Experimental Investigation. *The Quarterly Journal of Economics*, 108(2), 437–459.
- Fershtman, C., & Gneezy, U. (2001). Discrimination in a Segmented Society : An Experimental Approach Author (s): Chaim Fershtman and Uri Gneezy Published by: Oxford University Press Stable URL: <http://www.jstor.org/stable/2696452> JSTOR is a not-for-profit service that helps scholars , r. *The Quarterly Journal of Economics*, 116(1), 351–377.
- Grief, A. (1993). Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. *The American Economic Review*, 83(3), 525–548.
- Habyarimana, J., Humphreys, M., Posner, D. N., & Weinstein, J. M. (2007). Why Does Ethnic Diversity Undermine Public Goods Provision? *The American Political Science Review*, 101(4), 709–725. <https://doi.org/10.1017/S0003055407070499>
- Halevy, N., Bornstein, G., & Sagiv, L. (2008). “In-group love” and “out-group hate” as motives for individual participation in intergroup conflict: A new game paradigm: Research article. *Psychological Science*, 19(4), 405–411. <https://doi.org/10.1111/j.1467-9280.2008.02100.x>
- Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity an anonymous economic game. *Evolution and Human Behavior*, 26(3), 245–256. <https://doi.org/10.1016/j.evolhumbehav.2005.01.002>
- Henrich, J. (2000). Does culture matter in economic behavior ? Ultimatum Game Bargaining among the Machiguenga of the Peruvian Amazon by Does culture matter in economic behavior ? Ultimatum Game Bargaining among the Machiguenga of the Peruvian Amazon. *American Economic Review*, 4, 973–979. <https://doi.org/10.1257/aer.90.4.973>
- Henrich, J., Boyd, R., Bowles, S., Camerer, C. F., Fehr, E., Gintis, H., ... Tracer, D. (2005). “Economic man” in cross-cultural perspective: behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28(6), 795–815–55. <https://doi.org/10.1017/S0140525X05000142>
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of Homo economicus: Behavioral experiments in 15 small scale societies. *American Economic Review*, 91(2), 73–78. <https://doi.org/10.1257/aer.91.2.73>
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., ... Ziker, J. (2010). Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science*, 327(5972), 1480–1484. <https://doi.org/10.1126/science.1182238>
- Jakiela, P. (2011). Social Preferences and Fairness Norms as Informal Institutions: Experimental Evidence and numerous seminar participants for helpful comments. *American Economic Review*, 101(3), 509–513.
- Karlan, D. S. (2005). Using Experimental Economics to Measure Social Capital and Predict Financial Decisions. *The American Economic Review*, 95(5), 1688–1699. <https://doi.org/10.1257/000282805775014407>
- Kenzior, J., Zibika, J. P., & Voors, M. (2015). *Social relationships, local institutions, and the diffusion of improved variety seed and field management techniques in rural communities: six case studies in South Kivu, DRC.*
- Landa, J. (1981). A theory of the ethnically homogeneous middleman group: an institutional alternative to contract law. *The Journal of Legal Studies*, 10(2), 349–362.
- Larson, J. M. (n.d.). Networks and Interethnic Cooperation, (212).
- Leibbrandt, A., Ramalingam, A., Sääksvuori, L., & Walker, J. M. (2015). Incomplete Punishment Networks in Public Goods Games: Experimental Evidence. *Experimental Economics*, 18, 15–37. <https://doi.org/10.1007/s10683-014-9402-3>
- Lin, N. (1999). Building a Network Theory of Social Capital. *Connections*, 22(1), 28–51. <https://doi.org/10.1108/14691930410550381>
- Lin, N. (2002). *Social Capital: A theory of social structure and action* (Vol. 19). Cambridge University Press.
- List, J. A. (2005). The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions. *National Bureau of Economic Research, Working pa*, 1–51. <https://doi.org/10.1086/498587>
- Miguel, E., & Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in Kenya. *Journal of Public Economics*, 89(11–12), 2325–2368. <https://doi.org/10.1016/j.jpubeco.2004.09.004>
- Milinski, M. (2016). Reputation, a universal currency for human social interactions. *Philosophical Transactions of the Royal Society B*, 371(1687), 20150100. <https://doi.org/10.1098/rstb.2015.0100>
- Milinski, M., Semmann, D., & Krambeck, H.-J. (2002). Reputation helps solve the “tragedy of the commons.” *Nature*, 415(6870), 424–426. <https://doi.org/10.1038/415424a>
- North, D. (1997). *The Process of Economic Change. World Institute for Development Economics Research.* <https://doi.org/10.1093/acprof:oso/9780199242191.003.0002>
- Rockenbach, B., & Milinski, M. (2006). The efficient interaction of indirect reciprocity and costly punishment. *Nature*, 444(7120), 718–23. <https://doi.org/10.1038/nature05229>
- Sobel, J. (2002). Can We Trust Social Capital ? *Journal of Economic Literature*, 40(1), 139–154.
- Sommerfeld, R. D., Krambeck, H.-J., & Milinski, M. (2008). Multiple gossip statements and their effect on reputation and trustworthiness. *Proceedings of the Royal Society B: Biological Sciences*, 275(1650), 2529–2536. <https://doi.org/10.1098/rspb.2008.0762>
- Volk, S., Thöni, C., & Ruigrok, W. (2012). Temporal stability and psychological foundations of cooperation

- preferences. *Journal of Economic Behavior and Organization*, 81(2), 664–676.
<https://doi.org/10.1016/j.jebo.2011.10.006>
- Voors, M. J., Nillesen, E. E. M., Verwimp, P., Bulte, E. H., Lensink, R., & Van Soest, D. P. (2012). Violent conflict and behavior: a field experiment in Burundi. *The American Economic Review*, 102(2), 941–964.
- Wu, J., Balliet, D., & Lange, P. A. M. Van. (2016). Gossip Versus Punishment: The Efficiency of Reputation to Promote and Maintain Cooperation. *Scientific Reports*. <https://doi.org/10.1038/srep23919>

CHAPTER 6

Conclusion

“Economic action (like all action) is socially situated, and cannot be explained by individual motives alone; it is embedded in ongoing networks of personal relations rather than carried out by atomized actors”

~ Granovetter, (1992; p.4)

1 Overview

Global development tackles practical, but complex, challenges that reduce the quality of life of populations within developing countries. Global development practitioners attempt to implement solutions that bring about self-sustaining change and improve the well-being of those within underdeveloped regions. Such change ultimately remains rooted in individual actions. Whether it be targeting health with deworming pills or financial security through saving and lending opportunities – the impact of the solution is only catalysed by the action of individuals adopting them. The question of whether the solution “works” depends first on whether it is adopted and second on whether it generates the professed change in outcomes. As a result, program design must account for the forces that ultimately drive the actions and decisions of people in order to increase adoption of the innovations that can lead to improved wellbeing.

Throughout this dissertation, the term innovations has been broadly used, referring to both new information and new technologies. Adoption of an innovation is the decision to change one’s behaviour by putting into practice the new information or utilizing the new technology. Such decisions hinge on the belief that the change in behaviour will result in greater or equal well-being – termed ‘returns to adoption’. Only if an individual’s information set indicates that such returns are realizable is adoption likely to occur. This dissertation has focused on the ways in which social forces shape these information sets on innovations on which adoption decisions are based. Specifically it explores the relationship between the channels by which

individuals learn and agricultural productivity, diffusion of information through social networks, and the relationship between an individual's social preference and their network position.

Within many development initiatives there exists an implicit recognition of some of the social forces working behind individual decision-making. Particularly, the use of social learning to facilitate behavioural change is a popular strategy within development programs. The use of peer trainers is a popular tool within many development programs as a means of targeting knowledge gaps that may obstruct innovation adoptions. This is a particularly popular method within the fields of health (Kelly et al., 1991; Kim et al., 2015; Miguel & Gugerty, 2005; Oster & Thornton, 2012) and agriculture (Anderson & Feder, 2007; Conley & Udry, 2010; Kondylis, Mueller, & Zhu, 2014). Peer trainers act as the contact points into social networks from which information spreads from development practitioners to community members. Such designs rest on assumptions of social learning in which information on innovations is transmitted via channels of interpersonal relationships.

Agriculture has been a sector in which the use of peer trainers has become widespread and is most visible in the practice of agricultural extension programs (Anderson & Feder, 2004; Rivera, Quamar, & Crowder, 2001). Practitioners make use of farmer-trainers within communities in efforts to maximize program reach while minimizing direct costs to program implementers. With this approach, practitioners reduce the number of trainings to only farmer-trainers and then rely on social learning to spread extension information onward. Attention has been given to testing the overall effectiveness of agricultural extension on farmer productivity, income, and wellbeing (Anderson & Feder, 2004; Cunguara & Moder, 2011; Kondylis et al., 2014). However studies on the underlying mechanisms that shape social learning are only just beginning to emerge within the development economics literature (Banerjee et al., 2014; Beaman et al., 2013; Breza, 2016; Kim et al., 2015).

While development practitioners are eager to incorporate more effective methods into their program designs when possible, many lack the capacity or expertise to conduct rigorous empirical testing of mechanisms. Academics have the expertise necessary for formulating and empirically testing theories of how social forces shape and drive decisions. Through cooperation, a feedback loop is formed in which practitioners offer space for academic research to occur within their programs and in return receive insights on more effective

program design. Such partnerships have become very popular in testing the overall effectiveness of extension programs on metrics of technology adoption, productivity, and economic wellbeing (see Anderson & Feder, 2004 for a review of extension evaluations). However as data becomes richer and more accessible, opportunities are presented to test intermediary mechanisms that lay at the core of behavioural change and the adoption of new innovations. Studies have begun exploring the role of learning, social networks, and the diffusion of information through networks in attempts to better understand how social learning shapes adoption decisions (Beaman et al., 2015; Benyishay & Mobarak, 2015; Benyishay & South, 2013; Conley & Udry, 2010; Emerick, Janvry, Sadoulet, & Dar, 2016; Vasilaky & Leonard, *forthcoming*).

This thesis seeks to contribute to this growing empirical literature. First it explores tools of social learning and individual learning. Specifically it tests if reducing-costs to learning-by-doing complements social learning for greater impacts on farmer crop harvests (Chapter 2). Second, it dives into the social learning process by exploring diffusion of innovations through social networks. Two key questions are asked: (i) *who* participates in diffusion flows, (ii) *how far* do innovations diffuse, and (iii) how does (i) and (ii) vary by network entry-point (Chapter 3 and Chapter 4). Some development programs prioritize who benefits while others prioritise the reach of benefits. Different design elements (e.g. network entry point) may be better suited for one objective over another. Finally, the thesis culminates with a deeper dive into the relationship between an individual's social preferences and their social network position. As preferences lie at the heart of human decision-making, understanding the relationship between social networks and individual preferences can provide insight on the heterogeneous behaviour observed within communities.

This thesis is set within an agricultural context. Economic decisions are highly contextualized and the results of this thesis speak directly to decisions within agricultural production. However as more literature emerges on similar mechanisms in other sectors, patterns will emerge and mechanisms that are cross-cutting will become highlighted. Therefore this thesis is positioned as contributing to the overall literature of social learning and technology adoption, with a specific focus on agricultural innovations and rural communities. I review the major themes of the dissertation in the sections below. Within each section I discuss position insights from my research within a broader context and make recommendations for future research.

2 Channels of Learning

Individuals base their technology adoption decisions on the expected returns to adoption (Foster & Rosenzweig, 2010). These expectations are dynamic in that they are updated as relevant new information is received and incorporated. New information can be received through discussion with and observation of prior adopters (Chuang & Schechter, 2000; Nourani, 2016). However, some technologies have more heterogeneous returns or their proper application and structure of returns are less visible through social learning (Foster & Rosenzweig, 2010). Under such conditions, learning-by-doing can be a more effective channel of learning as data quality biases are reduced (Foster & Rosenzweig, 2010; Sunding et al., 2001). Learning-by-doing provides potential-adopters the opportunity to assess the marginal benefits to adoption given their personal situation. Within agricultural production, this means that farmers can experiment with new innovations in order to determine their respective marginal benefits given their production functions, market transaction costs, and optimization priorities (e.g. utility vs. profit). However experimentation comes at an investment cost, which can deter potential adopters.

In Chapter 2 of this thesis, an intervention that reduced the cost of learning-by-doing, coupled with a social learning program, is compared to a purely-social learning intervention. Results indicated that the combined intervention of social learning in conjunction with lower-cost experimentation had a greater impact on average farm output levels. This chapter speaks to how development programs can tailor designs for strengthened impact on wellbeing. Farmers are more strongly influenced by prior adopters that are comparable to themselves, but uncertainty still remains in regards to the quality and applicability of this information (Benyishay & Mobarak, 2015). Many agricultural technologies have a vital knowledge-component on proper use and application as well as heterogeneous returns depending on farmer application, soil quality, crop portfolio, and productive endowments. Targeting experiential learning channels in addition to social learning channels can strengthen overall outcomes on farmer output levels by providing farmers first-hand knowledge on their returns to adoption. However experimentation must be accompanied by sufficient information on the proper use of innovations. Without such information, negative effects could arise in which improper use yields negative returns, dampening adoption further.

Intuitively, one could anticipate that the synergies of the two learning channels increase over time. Farmers that adopt after experimentation will provide further observation points for farmers still deciding if adoption is the right choice for them (Nourani, 2016). These dynamic synergies could strengthen the cumulative impacts of programs in the longer run. The impact evaluation in Chapter 2 was conducted over a relatively short time-horizon for agricultural decisions, lasting only one year, so measuring such dynamic synergies is beyond the scope of this thesis. However future research is required to better understand potential dynamic synergies. Determining how such synergies play out within the timeline of innovation spread and what the optimal time is for introducing low-cost experimentation will be a key question for stronger program design.

Provision of subsidized samples of new innovations in facilitating learning-by-doing is costly. Alternative strategies for reducing the information quality bias may offer a lower cost. Research by BenYishay & Mobarak (2014) indicates that individuals weight information more heavily from those more comparable to themselves. Furthermore, Hanna et al. (2014) find that what information is noticed by potential-adopters when observing prior-adopters influences ultimate adoption decisions. Further research on program design elements that can ensure potential-adopters are receiving information from comparable, and trusted, sources as well as that the critical elements of information are being noticed could provide cost-effective strategies for increasing adoption rates.

3 Diffusion Targeting and Reach

Development programs vary in the extent to which they prioritize the targeting of their intervention towards specific beneficiaries compared to the ultimate reach of their intervention (Peyre Dutrey, 2007; Van Domelen, 2007). Program targeting decisions vary but some examples include goals of poverty alleviation (e.g. targeting most “in need”), by demographics (e.g. women or youth), or concerns of ‘elite capture’ (Alatas et al., 2013; Buvinic & Gupta, 1997; Platteau & Gaspart, 2003). Programs that prioritize reach are more focused on achieving widespread impacts. Neither program objective is inherently better or worse, but instead depends on the innovation in question and the capacity and mission of the organization. However targeting and reach are extremely pertinent to the study of social network diffusion. Program designs must take into account network diffusion dynamics and their relationship with outcomes of *who* is benefiting from the information flows as well as *how many* are benefiting.

Theoretical studies and empirical studies utilizing detailed online social network data have illustrated the importance of network structures and entry points on resulting diffusion patterns (Sinan Aral, Muchnik, & Sundararajan, 2013; Centola & Macy, 2007; Choi, Kim, & Lee, 2010). This thesis has extended this same question to the context of a developing country within the agricultural sector. Results from chapter 3 corroborate previous findings that in selecting network entry points, not all network measures create the same diffusion patterns. Studies show selecting network entry points that are more clustered together can have positive effects on diffusion reach (Beaman et al., 2015; Centola & Macy, 2007). However no consensus has been reached on the effectiveness of selecting individuals with the most connections as network entry points ((S. Aral & Walker, 2012; Gulyás & Imre, 2014; Hinz, Skiera, Barrot, & Becker, 2011; Kim et al., 2015) This thesis finds evidence that selection of network entry points by eigenvector centrality score yields little impact on the width of diffusion. However important results were found on who participated in diffusion processes as both central entry-points and isolate entry-points were found to transmit to network-central individuals (Chapter 3). This means that regardless of the centrality of the network entry-point, resources pool in the center of networks. This is an important finding especially as periphery households were found more likely to be female-headed, less wealthy, and less active in community leadership positions – all oft-used classifications of ‘marginal households’ that poverty alleviation programs may specifically target. Therefore, reaching marginal households may be better accomplished through strategies of direct targeting rather than through social diffusion. Additionally, in situations where innovations are more susceptible to elite capture, diffusion through social networks may compound potential adverse distributional effects.

While the results of Chapter 3 show no clear differentiation in reach of diffusion by centrality-selected entry points, this does not mean that other network parameters will have no effect. Each measure of network connectedness capture different relationships between information flows and the network structure. Further research on how these relationships vary by entry-point selection processes using different network metrics will allow these relationships to be leveraged within program designs. Results from Chapter 3 indicate that eigenvector centrality is not an ideal entry-point selection parameter for the agricultural innovation of in-organic fertilizer. The knowledge-intensity of in-organic fertilizer may require greater consideration be given

to the proximity of the entry-points to one another in order for potential adopters to be sufficiently exposed so as to be swayed into adopting (as seen in Beaman et al., 2015).

Chapter 4 uses a complex contagion simulation to network structures, which accounts for proximity of network-entry points in predicting which pair of entry-points maximizes diffusion reach. Diffusion reach is defined as the proportion of the community who ‘adopted’ the innovation within the simulation. Use of the optimal entry-points results in diffusion rates of over 50% within the simulation. The high costs of collecting network data and the strict cost constraints that development practitioners face may make network simulation-based selection processes impractical. An alternative, lower-cost strategy of community-wide polling is therefore compared within the simulation. While community-wide polling is not found to be perfectly correlated with simulation-identified optimal entry points, there is substantial overlap. Our results are in line to previous findings on community-based selection by Valente & Davis (1999). Community-polling may offer an opportunity to leverage network diffusion dynamics without the high costs of network-data. Field testing of this approach is required before community-wide polling can definitively be put forward as an effective alternative to network-based selection. A study by Emerick et al. (2016) finds that selection of network contact-points using a community-wide meeting yields insignificant results on diffusion. However, community meetings are potentially susceptible to hierarchical social pressures that do not give equal voice to all attendees’ individual knowledge of network structures. Unequal representation may result in less complete network information behind entry-point selection. As a result, field testing a private community-polling result would be beneficial in highlighting both the effectiveness of community-wide selection on diffusion as well as shedding light on dynamics of public versus private nomination processes.

4 Networks and Individuals

A third dimension of social networks that cross-cut the chapters of this thesis is how social networks and individual characteristics are related. Chapters 3 and 4 both find evidence of an individual’s centrality level being correlated with socioeconomic indicators of wealth, education, community leadership, and gender. These results suggest that network centrality is predictive of social prominence. Furthermore, Chapter 4 finds evidence of correlations between peer-perceptions of influence with network centrality. This strong

relationship may offer a policy opportunity. Selection using socio-economic indicators offers a strong correlation to both network connectedness and perceived social influence. This predictive relationship offers development practitioners a lower-cost method of identifying entry-points that leverage diffusion dynamics and social influence.

Social influence, as measured by community nomination, is not perfectly correlated with either network centrality nor socio-economic prominence. This indicates that there are missing explanatory factors of influence within the model. One explanation may lie in individual preferences. Socially influential individuals were found to be more pro-social in terms of trust, trustworthiness, and cooperation relative to their non-nominated peers (Chapter 5). Social preferences are important for understanding which individuals are more likely to participate in community-beneficial interventions, even at individual costs, and how such individuals can be incentivized. For development programs, this information can be leveraged within program design for selecting more pro-social individuals and designing more effective incentives.

More research is necessary on what incentive structures will best facilitate greater diffusion. Evidence suggests that social learning can be strengthened when communicators are provided with small incentives (Benyishay & Mobarak, 2015). Exploring how response to different incentive structures varies by social preference sets and network position will help identify how incentives can be better used for generating diffusion based on network entry-point.

This thesis explores only preferences as they relate to sociality. Exploring relationships between network position with a broader set of preferences (e.g. risk, time-discounting) could offer greater insights behind network position and innovation adoption decisions. Preference under-gird economic decision-making. Understanding links between an individual's likelihood to adopt an innovation, preference set, and network position will offer opportunities for identifying pro-social first-adopters that are well positioned for catalysing diffusion.

5 Final Thoughts

Social networks, and the interpersonal relationships that define them, can provide valuable information on how global development initiatives can strengthen the spread of innovations. However the complex

relationships between technology adoption, social influence, and networks within developing countries is still in the early stages of research. It is on the shoulders of academics to more thoroughly explore the mechanisms that underlie these relationships so they can be translated into policy designs that increase intervention impacts.

Opportunities for such research continue to grow as data costs decrease with the continued expansion of ICT infrastructure within developing countries. This data will provide opportunities for researchers to explore more complex and contextualized questions on the relationships between social networks and human behaviour. This zooming in on the foundations of human behaviour is imperative in understanding how social forces drive the individual decisions that ultimately shape development program outcomes.

This thesis was directed by zooming in on mechanisms, but the digitalization of social interactions will create opportunities to explore networks at a regional, national, and international level. Network data within this thesis was collected through in-person surveys which necessitated that only small networks, defined at the community level, be analysed. Future research is required on larger community networks as it is unlikely that network dynamics scale up in a linear fashion. Furthermore, future research should not ignore networks of communities and the interlinkages between each community. By moving up a level and exploring networks of communities, a better understanding of geo-spatial diffusion within regions becomes possible. Such insights could guide decisions on which communities should be targeted by development programs as opposed to just which individuals within those communities.

References

- Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., Purnamasari, R., & Wai-poi, M. (2013). Does Elite Capture Matter? Local Elites and Targeted Welfare Programs in Indonesia. *NBER Working Paper*, (18798). <https://doi.org/10.3386/w18798>
- Anderson, J. R., & Feder, G. (2004). Agricultural extension: Good intentions and hard realities. *World Bank Research Observer*, 19(1), 41–60. <https://doi.org/10.1093/wbro/lkh013>
- Anderson, J. R., & Feder, G. (2007). Chapter 44 Agricultural Extension. *Handbook of Agricultural Economics*, 3(6), 2343–2378. [https://doi.org/10.1016/S1574-0072\(06\)03044-1](https://doi.org/10.1016/S1574-0072(06)03044-1)
- Aral, S., Muchnik, L. E. V., & Sundararajan, A. (2013). Engineering social contagions: Optimal network seeding in the presence of homophily. *Network Science*, 1(2), 125–153. <https://doi.org/10.1017/nws.2013.6>
- Aral, S., & Walker, D. (2012). Identifying Influential and Susceptible Members of Social Networks. *Science*, 337, 337–341. <https://doi.org/10.1126/science.1215842>
- Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2014). *Gossip: Identifying Central Individuals in a Social Network* (NBER Working Paper Series No. 20422). National Bureau of Economic Research. <https://doi.org/10.3386/w20422>
- Beaman, L. a, Udry, C. R., Karlan, D., & Thuysbaert, B. (2013). Profitability of Fertiliser: Experimental Evidence from Female Rice Farmers in Mali. *The American Economic Review*, 103(3), 381–386. <https://doi.org/10.1257/aer.103.3.381>
- Beaman, L., Benyishay, A., William, C., & Mobarak, A. M. (2015). *Can Network Theory based Targeting Increase Technology Adoption ?*
- Benyishay, A., & Mobarak, A. M. (2015). *Social Learning and Incentives for Experimentation and Communication*.
- BenYishay, A., & Mobarak, A. M. (2014). Social Learning and Communication. *NBER Working Paper*, (20139). <https://doi.org/10.3386/w20139>
- Benyishay, A., & South, N. (2013). Communicating with Farmers through Social Networks †.
- Breza, E. (2016). *Field Experiments, Social Networks, and Development. Working Paper*.
- Buvinic, M., & Gupta, G. R. (1997). Female-headed households and female-maintained families: are they worth targeting to reduce poverty in developing countries? *Economic Development and Cultural Change*, 45(2), 259–280.
- Centola, D. M., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3), 702–734. <https://doi.org/10.1086/521848>
- Choi, H., Kim, S.-H., & Lee, J. (2010). Role of network structure and network effects in diffusion of innovations. *Industrial Marketing Management*, 39(1), 170–177. <https://doi.org/10.1016/j.indmarman.2008.08.006>
- Chuang, Y., & Schechter, L. (2000). Social Networks in Developing Countries. *Annual Review of Resource Economics*, (7). <https://doi.org/10.1146/annurev-resource-100814-125123>
- Conley, B. T. G., & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *The American Economic Review*, 100(1), 35–69.
- Cunguara, B., & Moder, K. (2011). Is Agricultural Extension Helping the Poor? Evidence from Rural Mozambique. *Journal of African Economics*, 20(4), 562–595.
- Emerick, K., Janvry, A. De, Sadoulet, E., & Dar, M. H. (2016). *Identifying early adopters , enhancing learning , and the diffusion of agricultural technology*.
- Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of Technology Adoption. *Annual Review of Economics*, 2. <https://doi.org/10.1146/annurev.economics.102308.124433>
- Granovetter, M. S. (1992). Economic Institutions as Social Constructions : A Framework for Analysis. *Acta Sociologica*, 35(1), 3–11.
- Gulyás, G. G., & Imre, S. (2014). Measuring Importance of Seeding for Structural De-anonymization Attacks in Social Networks. In *Pervasive Computing and Communications Workshop (PERCOM), IEEE International Conference*.
- Hanna, R., Schwartzstein, J., Benjamin, D., Gentzkow, M., Halac, M., Newman, A., & Ogden, T. (2014). Learning Through Noticing: Theory and Experimental Evidence in Farming.
- Hinz, O., Skiera, B., Barrot, C., & Becker, J. U. (2011). Social Contagion – An Empirical Comparison of Seeding Strategies for Viral Marketing. *Journal of Marketing*, 75(6), 55–71. <https://doi.org/http://dx.doi.org/10.1509/jm.10.0088>
- Kelly, J. A., St. Lawrence, J. S., Diaz, Y. E., Stevenson, L. Y., Hauth, A. C., Brasfield, T. L., ... Andrew, M. E. (1991). HIV risk behavior reduction following intervention with key opinion leaders of population: An experimental analysis. *American Journal of Public Health*, 81(2), 168–171. <https://doi.org/10.2105/AJPH.81.2.168>
- Kim, D. A., Hwong, A. R., Stafford, D., Hughes, D. A., O'Malley, A. J., Fowler, J. H., & Christakis, N. A. (2015). Social network targeting to maximise population behaviour change: A cluster randomised controlled trial. *The Lancet*, 386, 145–153. [https://doi.org/10.1016/S0140-6736\(15\)60095-2](https://doi.org/10.1016/S0140-6736(15)60095-2)
- Kondylis, F., Mueller, V., & Zhu, S. J. (2014). *Seeing is believing? Evidence from an extension network experiment* (Policy Research Working Paper No. 7000).
- Miguel, E., & Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in Kenya. *Journal of Public Economics*, 89(11–12), 2325–2368. <https://doi.org/10.1016/j.jpubeco.2004.09.004>

- Nourani, V. (2016). *Social Network Effects of Technology Adoption : Investing with Family , Learning from Friends & Reacting to Acquaintances*.
- Oster, E., & Thornton, R. (2012). Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-Up. *Journal of the European Economic Association*, 10(6), 1263–1293. <https://doi.org/10.1111/j.1542-4774.2012.01090.x>
- Peyre Dutrey, A. (2007). *Successful Targeting? Reporting Efficiency and Costs in Targeted Poverty Alleviation Programmes*. *Social Policy and Development Programme Paper* (Vol. 35). Retrieved from [http://www.unrisd.org/80256B3C005BCCF9/\(httpAuxPages\)/0B87C67449C938EDC12573D10049830B/\\$file/Peyrepap.pdf](http://www.unrisd.org/80256B3C005BCCF9/(httpAuxPages)/0B87C67449C938EDC12573D10049830B/$file/Peyrepap.pdf)
- Platteau, J., & Gaspart, F. (2003). The Risk of Resource Misappropriation in Community- Driven Development. *World Development*, 31(10), 1687–1703.
- Rivera, W. M., Quamar, M. K., & Crowder, L. V. (2001). *Agricultural and rural extension worldwide: options for institutional reform in the developing countries* (Sustainable Development Department). Rome.
- Sunding, D., Zilberman, D., Bruce, L. G., & Gordon, C. R. (2001). Chapter 4 The agricultural innovation process: Research and technology adoption in a changing agricultural sector. *Handbook of Agricultural Economics, Volume 1*, 207–261. [https://doi.org/http://dx.doi.org/10.1016/S1574-0072\(01\)10007-1](https://doi.org/http://dx.doi.org/10.1016/S1574-0072(01)10007-1)
- Valente, T. W., & Davis, R. L. (1999). Accelerating the Diffusion of Innovations using Opinion Leaders. *The ANNALS of the American Academy of Political and Social Science*. <https://doi.org/10.1177/0002716299566001005>
- Van Domelen, J. (2007). *Reaching the Poor and Vulnerable: Targeting Strategies for Social Funds and other Community-Driven Programs* (No. 711). *Social Protection Discussion Paper*. Retrieved from <http://siteresources.worldbank.org/SOCIALPROTECTION/Resources/SP-Discussion-papers/Social-Funds-DP/0711.pdf>
- Vasilaky, K., & Leonard, K. (n.d.). As Good as the Networks They Keep? Improving Outcomes through Weak Ties in Rural Uganda. *Economic Development and Cultural Chang*.

ENGLISH SUMMARY

This thesis contributes to a growing literature that explores relationships between social networks and innovation diffusion within a developing country context. Given this context, the networks of interest within this thesis are the offline interpersonal relationships between community members. Diffusion channels for new innovation are therefore limited to word-of-mouth communication, observation, and personal experience.

Chapter 2 of this thesis analyses two policy tools in targeting these information gaps. The first is through social learning as part of a farmer extension program. The second combines social learning with experiential learning, reducing the cost to personal experimentation with subsidized improved input packages. Our results indicate that farmers who are exposed to both social learning and learning-by-doing more significantly impacts farmer productivity relative to those receiving no intervention and those exposed only to social learning. I interpret this result as an indication of learning-by-doing combined with social learning being a more effective strategy for facilitating adoption of technologies that have more heterogeneous returns to adoption.

Chapter 3 of this thesis tests the difference in diffusion patterns that result by varying the network contact-point. Specifically, network contact-points are selected as being either the most central or least central individuals within the network. I find evidence that centrality affects the speed of distribution but does not affect the width of diffusion nor which individuals are participating within the diffusion process. Furthermore, large attenuation is observed throughout the diffusion process, which suggests the importance of selecting a sufficiently large set of lead community members for the spread of new technology.

Chapter 4 combines a community-wide polling of network entry-points combined with detailed community network and socio-economic data. First we explore what attributes are prioritized by community members in nominating a resident farmer as an extension contact-point. Second, we use simulations to compare the diffusion spread of top-nominated individuals as network entry-points compared to entry-points that achieve maximal spread within diffusion simulations. We find that community members prioritize network connectedness, pro-social preferences, and socioeconomic indicators of gender, age, formal leadership, and education levels within their nomination decisions. Furthermore, receiving the top three most amount of

nominations is found to be significantly correlated with selection as an optimal entry-point within the diffusion simulation. These results suggest that community-wide polling offers a less data-intensive opportunity to realize gains in diffusion warranted through network-based seeding.

Chapter 5 explore whether an individual's observed social preferences is correlated with an individual's centrality within the network structure. Our results indicate that individuals with high centrality are more trusting and more trustworthy than individuals with lower centrality. Moreover, individuals with low centrality are treated worse in these interactions—people trust them less initially, and return less money to them. Within a group context, little evidence is found of more central individuals displaying more cooperative behavior. Instead, for group cooperation, when a single monitor can observe contribution decisions, the presence of a direct link and more mutual network connections with a monitor correlates to more cooperative behavior by that individual. Our results suggest that network centrality and pro-social preferences are related but more localized network ties are more strongly correlated with pro-sociality than overall network connectedness.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor, Maarten Voors. It has been a privilege to be his first PhD student. Without his guidance and encouragement this PhD would not have been achievable. Through his time, ideas, enthusiasm and good example he has taught me how good research is conducted. Maarten's infectious passion for his work provided motivation at even the most challenging points of my PhD. Having had no prior experience in field research within Africa, my learning curve was a steep one and challenges were plenty. Maarten struck the right balance of providing guidance and support while allowing me the space to carve my own path and grow into my own version of a researcher. This careful balance of *laissez-faire* with a strong safety net carried through each progressive phase of my PhD, allowing me to grow not only as an economist but as a leader, a team member, and a person.

Special thanks also to Erwin Bulte who saw in me potential that I could not yet see and for giving me the opportunity to join the Development Economics Group at Wageningen. I greatly appreciate that his door was always open and advice readily available to help refine and strengthen ideas at all stages of research. It has been a privilege working under his supervision and I am grateful to have had the opportunity to observe and learn from his example.

I would also like to thank all my colleagues at DEC, both past and present, who contributed immensely to both my professional and personal life at Wageningen. Within this group I have benefited from thoughtful advice, feedback, guidance, and friendships. A special thank you also to all those I have had the good fortune of collaborating on projects and papers with, in particular Peter van der Windt, Koen Leuveld, Janneke Pieters, and Lonneke Nillesen. Through these collaborative efforts I benefitted from the experience, feedback, and support of each and grew into a better researcher as a result.

This PhD would not have been possible without the hard work in the field collecting the data of Koen Leuveld, John Quattrochi, and Paul Hofman. These efforts were fruitful only as a result of the dedication of their research assistants and the cooperation and support of the study participants. This PhD is a reflection of the hard work, time, and efforts invested by each of these groups and I am grateful for being able to use the resulting data within this thesis.

Furthermore, I cannot overlook the research team and participants of the early projects I worked on in Sierra Leone. While none of the resulting data features in this thesis, the experience on those projects shaped who I am as a researcher. I am especially grateful to those colleagues who kept me sane during the long and challenging hours that field research demand: Thijs van Bommel, Karen van Zaal, Paul Hofman, Esther Mokuwa, Lizzy van der Wal, and Frazer Sinclair. I was fortunate to learn the ropes of field work with an amazing team of research assistants and appreciate their patience and support as I grew into the role as a research field manager. While the entire team is too large to personally list everyone, I value the contributions of each team member and their dedication to the research despite long and gruelling hours in difficult conditions. I have incredible respect for all that they were able to accomplish and admiration for the enthusiasm with which they undertook the work.

While I have thanks many individuals for all their professional contributions and support, I cannot overlook how much I have enjoyed their friendships. I am grateful for all the coffee breaks, after-work beers, and moral support that made my PhD experience so enjoyable. A special thank you to Koen Leuveld, Aster Postma, and Heart Patient, Negra, and Louwe for welcoming me (and Pixel) into their home and offering their guest bedroom whenever I needed to commute up from Brussels. No amount of Belgian beer and chocolate will ever make up for your generosity and all the home cooked meals.

Lastly, I would like to thank my family for all their love and encouragement. To my parents for raising me to have the independence and the courage to face new experiences, learn from my failures, and pursue my

ambitions – I needed each of these in abundance throughout this PhD. To my sisters for their visits, calls, e-mails, and encouragement, I am so fortunate to have such a strong support network. And lastly to Jason, for never holding against me my frequent absences, incoherent ramblings, passionate vent sessions, or periodic cycles between procrastination and frantic productivity – I could not have done this without your love and support.

Martha Ross
Wageningen School of Social Sciences (WASS)
and PE&RC
Completed Training and Supervision Plan



Name of the learning activity	Department/Institute	Year	ECTS*
A) Project related competences			
Advanced Macroeconomics, ECH30806	WUR	2012	6
Economics of Networks	Tinbergen Institute	2014	3
Applied Micro Econometrics II: Empirical Treatment Evaluation	Tinbergen Institute	2014	3
Research Proposal writing	WASS	2013	6
ArcGIS 10.1 Introduction	WUR GeoDesk	2013	1
B) General research related competences			
Research Methodology: From Topic to Proposal	WASS	2012	4
Introduction course	WASS	2012	1
Techniques for Writing and Presenting a Scientific Paper	WGS	2015	1.2
<i>'Bargaining Markets, Information Asymmetry, and Social Capital'</i>	CSAE, Oxford, UK	2015	1
<i>'Lessons from the Field'</i>	EGAP Learning Days, Accra, Ghana	2015	1
<i>'A Farmer's Knowledge and Money: Testing Constraints to Closing Smallholder Yield Gaps in Eastern DRC'</i>	DEC Seminar Series	2015	1
<i>'Going beyond the fields understanding Yields Gaps in Eastern DRC'</i>	Workshop for Research on Kivu Region of DRC, Antwerp, Belgium	2015	1
C) Career related competences/personal development			
Entrepreneurship in/outside Science	StartLife	2015-2016	1.6
Training Field Staff on Research Protocols for 5 Research Projects in Sierra Leone and Republic of Congo	WUR and research partners	2012-2015	6
Course Teaching Assistant: Cost-Benefit Analysis	WUR	2014	0.4
Course Teaching Assistant: Methodology for Field Research in the Social Sciences	WUR	2015	0.2
Total			37.4

*One credit according to ECTS is on average equivalent to 28 hours of study load

The research described in this thesis was made possible through the financial support of the Economic and Social Research Council (ESRC) Growth Research Programme (grant ES/J009008/1) and the Wageningen IPOP Programme - Sustainable and Smart Food Supply .