# How do Farmers Learn from Extension Services? Evidence from Malawi

Annemie Maertens, Hope Michelson and Vesall Nourani*

25 November 2018

**Abstract**

Agricultural extension services can play an important role in increasing farmer yields and incomes yet evidence of the effectiveness of extension services in Sub-Saharan Africa has been mixed. We study farmers' learning about agricultural technologies using a (quasi) randomized controlled trial in which farmers differ in their exposure to commonly used extension methods that range in their intensity of interaction. We find that farmers who participated in season-long farmer-led demonstration plot cultivation learn about the critical adoption details of production processes and adopt more components of a new, multi-component technology. Farmers invited to attend farmer field day events learn considerably less about production process details. Building on qualitative interviews, we then develop a two-stage learning process in which farmers first form yield expectations and then choose how much to invest in learning the details of the production processes subject to yield beliefs and the learning costs. We test this model using detailed data on beliefs, knowledge, adoption and constraints and find evidence that farmers' yield beliefs hinge around observed yield, and these observed yields affect learning efforts.

*Annemie Maertens (a.maertens@sussex.ac.uk) at Sussex University; Hope Michelson (hopecm@illinois.edu) at the University of Illinois at Urbana-Champaign and Vesall Nourani (vnourani@mit.edu) at MIT. We thank the field team of Wadondo Consulting for collecting data and Eric Kaima and Ephraim Chirwa for supervising the data collection. We acknowledge excellent research support from Tsegay Teklelassie and Kwabena Krah. We are grateful to Chris Barrett, Brad Barham, Ephraim Chirwa, Cheryl Palm, Wezi Mhango, Matt Embrey, J.V. Meenakshi, Giordano Paloni, and seminar participants at the AAEA meetings 2017, NEUDC 2017, 3IE-Delhi, 3IE-London and AGRA-Nairobi, the University of Gottingen for excellent comments and suggestions. We acknowledge financial support from the International Initiative for Impact Evaluation under TW4.

# Introduction

Agricultural extension can play a crucial role in overcoming farmer information constraints and encouraging the adoption of improved agricultural technologies, thereby increasing yields and incomes (Birkhauser et al. 1991, Picciotto and Anderson 1997, Feder et al. 1999, Anderson and Feder 2007, Davis 2008, Birner et al. 2009, Waddington and White 2014). Supporting and enhancing cost-effective agricultural extension systems is especially important in developing countries where the economy centers on agriculture. In Malawi, where this study is situated, agriculture represents 30 percent of the GDP and employs nearly 70 percent of the population (2013 World Development Indicators). As information is commonly non-rival and non-excludable, primary responsibility for developing and disseminating agricultural education programs falls to governments, especially at the start of diffusion processes where private sector operations are unable or unwilling to perform these services.[1] Where responsibility resides with government extension services, effectiveness can vary. Birkhauser et al. (1991) reports rates of return ranging from negative to over 100 percent (see also IEG 2011, Anderson and Feder 2007 and for an overview of evaluations set in Malawi: Ragasa and Mazunda 2018).

A share of the variability in extension effectiveness may be attributable to the range of extension models employed, from systems in which extension agents visit and train lead farmers, to farmer-led demonstration plots where farmers experiment with a new technology under the guidance of an extension agent, and field-days in which farmers learn about technologies on-site for a single day (see Anderson et al. 2006). Such extension models not only range widely in terms of time and expense to both farmers and implementing agencies, but also have different implications for farmer learning and hence might result in different adoption patterns.

Demonstration plots, for instance, are often sited in the same village where participant farmers live, i.e., where soil and climatic conditions are likely to be familiar and similar to most participants' conditions. Field-days, in contrast, often take place further from farmers' own communities, where local conditions might be quite different (and perhaps unknown) than those faced by visiting farmers at home. As returns to agricultural technologies are heterogeneous and depend on these conditions (Marenya and Barrett 2009, Dulfo et al. 2008, Suri 2011), farmers are more likely to learn something useful about the profitability of a new technology when the demonstrations are situated nearby.[2]

[1] Krishnan and Patnam (2012), for instance, note that in Ethiopia, in 1999, government extension agents served as farmers' primary source of information about hybrid seeds, but that ten years later private-sector seed companies had almost entirely assumed this role.

[2] This is, assuming that farmers learn locally, they will learn locally, i.e., contingent upon soil, climate,

In addition to learning about profits, farmers learn about enhancements to the production process including the optimal use of inputs. In Conley and Udry (2010), Ghanian farmers concentrate on one dimension of a technology: the optimal amount of fertilizer on pineapple, a new crop in the region. However, other critical sets of technologies have proved more difficult to promote in Sub-Saharan Africa. These technologies typically involve adjusting among, and hence learning about, multiple production dimensions (see Beaman et al. 2013, Emerick et al. 2016 and Bulte et al. 2014 and Mponela et al. 2016). In this study, we focus on one of these multi-dimensional technologies: Integrated Soil Fertility Management Practices (henceforward ISFM), a group of techniques designed to increase the fertility of soils. ISFM includes application of mineral fertilizers, incorporation of organic matter, adoption of agroforestry, crop rotation and intercropping with legumes, and use of conservation agriculture practices. Learning about these numerous dimensions can be demanding and farmers make choices about what to pay attention to (see Schilback et al. 2016 and Lichand and Mani 2017).[3] Hanna et al. (2014), for instance, show that seaweed farmers in Indonesia only learn to be attentive to pod size – an important input dimension – after being presented with simple information pointing out that it is critical. The important practical implication of this work: extension service models which include learning-by-doing and repetition, such as demonstration plots, might be effective because they reduce the cognitive demands on farmers, and result in better retention of knowledge (on the value of repetition in learning, see Brown et al. 2014, VanLehn 1996, Kim et al. 2013).

If learning about the production process comes at a significant cognitive cost, farmers might engage in what has been termed "rational inattention" (see Ghosh 2016 for a theoretical approach on rational inattention; and Gabaix 2017 for an overview). This might imply a two-stage learning process, as in Nourani (2017), where farmers first establish a belief of the profitability, and only when this belief meets a particular cut-off or set of conditions, proceed with learning the production processes. Moreover, farmers using attention strategically may focus on dimensions of the technology where the perceived benefits might be most likely to exceed the perceived costs. For instance, credit-constrained farmers might focus on the more labor-intensive dimensions of a new production process.

The contribution of this paper is to exploit an opportunity that introduced variation into

etc.; only a yield-draw from a plot which shares these conditions is likely to impact their own thinking and operations (on the implications of heterogeneity for social learning, see, among others, Munshi 2004, Tjernstrom 2016 and Crane-Droesch 2018; on adherence to Bayesian learning, see, among others, Lybbert et al. 2007).

[3] See Kahneman (1973), Kahneman (2003), Gabaix et al. (2006), Fehr and Rangel (2011), Harstad and Setlen (2013), and Rabin (2013) for an introduction to bounded rationality models.

farmer exposure to different extension methods instructing farmers about a single (complex) agricultural technology. We study the effects of farmer field-days and farmer demonstration plots on farmer learning and adoption and we present a model of farmer learning based on the insights from the evaluation. We exploit rich data on soil conditions, demonstration plot performance, and agronomic outcomes to understand how farmers direct their attention under time constraints and how spatial variability in growing conditions impacts farmer learning and adoption. To be clear at the outset: this paper is not a horserace between extension models. Instead, we are interested in contrasting field-days and demonstration plots to gain broader and more generalizable insights into farmer learning.

The literature on the effects of extension struggles with two primary empirical challenges. First, farmers who seek out and receive extension services might be more skilled and motivated than farmers who do not seek such services.[4] Moreover, areas that attract extension services are also often areas with better agronomic potential. Because such factors are often unobserved by researchers, they can cause omitted variable bias, threatening the causal interpretation of estimated parameters. A second challenge is that although an extension program may be successful in terms of knowledge diffusion, adoption among farmers may be influenced by other factors (market failures, logistical challenges, etc.); and learning may not always translate into adoption. As standard household surveys often do not detail the learning process, studies have often faced challenges discerning whether such failures reside in the education process itself, or in other circumstances down the line.

We designed our study to meet these challenges. We worked in partnership with the Clinton Development Initiative (henceforward CDI) in Malawi. CDI has set up a program of farmer-led demonstration plots and field-days aiming to disseminate information about ISFM - with a focus on the maize and soybean cropping systems. We selected 250 villages and randomized access to CDI's field-days. This process eliminates biases originating from between-village unobserved variation when establishing the effects of access to field-days. Using detailed village-level data we note that the villages in which CDI placed demonstration plots (which, unlike the field-days, were not randomized) were comparable to the villages which merely received access to the field-days. This observation allows us to use a regression approach when establishing the impacts of demonstration plots. We complement this quasi-randomized design with a household panel survey documenting the adoption of ISFM technologies, as well as farmer knowledge of ISFM technology processes and yield expectations.

We begin our analysis by establishing the impacts of demonstration plots and field-days. We find that farmers who participate in demonstration plots plan to adopt around, on

[4] See, for instance, Owens et al. (2003) who document the extent of such a bias in Zimbabwe.

average, 22 percent more of the recommended ISFM technologies one year after the program's start compared to similar farmers in control villages. Farmers who participate in a field-day, on the other hand, do not plan to adopt more of the ISFM technologies relative to similar farmers in the control villages. This lack of adoption might be due to a lack of learning: One year after the program was introduced, farmers who participate in demonstration plots score, on average, 8 percent higher on a test measuring knowledge of ISFM, compared to similar farmers in control villages. Farmers who attend a field-day do not score any higher compared to similar farmers in control villages.

To gain a better understanding of the learning process, we conducted focus group interviews. Farmers who attended field-days report being impressed by the yields on the field-day plots they visited, and reported being convinced that pesticides, and more generally, "modern inputs" are important. However, when we inquired for instance about specifics related to pesticide use, few of the field-day participants knew brand names, where to purchase these inputs, or how to prepare and apply the products. In other words, though farmers did acquire some knowledge about improving yields, they retained little specific information about production practices. Farmers who participated in a demonstration plot, on the other hand, tended to recall details about the inputs and production practices. For instance, these farmers were able to recount the inoculation process for soybean seeds, from preparing the inoculant to covering the seeds and planting them on ridges.

This narrative is consistent with the presence of a rational, but costly, two-step learning process in farmers first assess profitability, and only then invest in learning subject to constraints. Field-day participants, by virtue of being matched to a plot which is further afield, learn less about the profitability of the technology, and being less convinced about the benefits of the technology, as well as facing a high cognitive cost of learning (due to time constraints), focus their attention on only certain dimensions of the technology. Demonstration plot participants, on the other hand, learn substantially more about the profitability of a given technology, and face fewer cognitive (and time) constraints. Hence, conditional on what they learn about the profitability, they can come away with stronger comprehension of multiple dimensions of a complex technology process.

We present a learning model based on these insights and test its implications in our data.

First, the model predicts that yield expectations should respond to observed yields, and respond more so if the conditions of the extension demonstrations more closely match the farmer's own conditions. We test this prediction in the demonstration plot villages where we collected and analyzed soil samples from the demonstration plot and the farmers; as well as demonstration plot yield data. We find that the farmers' yield expectations correlate more strongly with the observed yields of soybean and maize if the farmers' soil is more similar to the demonstration plot's soil.

Second, this model predicts that the amount of cognitive effort farmers commit to learning the production process responds to these yield expectations. We again test this prediction in the demonstration plot villages using rainfall data. We note a positive correlation between demonstration plot yields and the farmers' knowledge of ISFM production practices. As this correlation could be attributed to reverse causality: demonstration plot farmers with increased knowledge more accurately followed CDI's guidelines, and hence, these demonstration plots had a higher yield; we repeat the regression analysis using the rainfall data instead of the yields. This analysis builds on the observation that germination rate, i.e., the percentage of seeds which germinate about 3 weeks after planting, correlate strongly with yields, but are largely determined by rainfall patterns at the start of the season. We confirm that a later start of the rains, and additional flood days (days with more than 50 mm/day), negatively correlate with the knowledge farmers hold about the ISFM processes (this result holds for soybean only).

Third, the model notes the dependency of the learning process on wealth and cognitive costs; as farmers who are constrained in terms of time and cognitive resources focus on the technologies they can realistically adopt. We find that demonstration plot participants' knowledge of ISFM practices for capital-intensive technologies, such as inputs for soybean cultivation, indeed positive correlate with wealth and observed demonstration plot yields. For farmers who were invited to participate in field-days and who arguably face a significantly higher cognitive cost of learning, we note a similar correlation between being credit constrained and learning about soybean cultivation (it is notable that we find no such result for hybrid maize, for which recommended technologies are more labor rather than credit-intensive). This is consistent with the accounts of the field-day participant who noted that, because they were credit-constrained, they preferred to focus on learning about (labor-intensive) technologies for maize, such as mulching, and optimal planting practices.

These results give a new meaning to the "rational but poor" farmers thesis originally proposed by Schultz (1964) and, subsequently tested by many, among others, Hopper (1965). Indeed, while farmers in our study might appear "irrational" at first in that they show evidence of inattention to technologies which might be beneficial, once one takes into account constraints imposed by heterogeneity and cognitive resources, the learning process appears rational. In effect, farmers in our study appear to decide how much effort to put into a learning experience, and to focus on specific aspects which they find important. These aspects are determined by individual farmers' expectations and constraints including market constraints.

The structure of the rest of this paper is as follows. In the next section, we introduce CDI's program. In the third section we present the study design, data collected, and descriptive statistics. The fourth section presents the results of the impact evaluation component: the effects of demonstration plots and field-days on adoption and knowledge. In section five we present a model of rational inattention, of which we test the implications in section six. Section seven concludes with a discussion on the implications for extension systems in Sub-Saharan Africa.

# CDI's ISFM extension program

ISFM includes a range of agricultural technologies to improve the use of nutrients and water and to increase crop yields. These practices include the combined use of mineral fertilizers, soil amendments (such as lime and rock phosphate) and organic matter (crop residues, compost and green manure); agroforestry (the combination of crops and trees including Nitrogen fixing "fertilizer trees"[5]), crop rotation and intercropping with legumes and conservation agriculture (no-till farming that uses a combination of mulch, direct planting and crop rotation to maintain fertility, prevent erosion and suppress weeds) (cited from AGRA and IIRC 2014). Thus, by improving the health of the soil, i.e., its ability to store and gradually release nutrients and water, ISFM, both directly and indirectly improves yields through increasing effectiveness of other inputs (Marenya and Barrett 2007).

Soil fertility is low and declining in Sub-Saharan Africa (see, among others, Tully et al. 2015 and Njoloma et al. 2016, Sanchez 2002). Hence, the benefits of ISFM in terms of increasing average yields and reducing yield variance can be substantial (see, among others, Kerr et al. 2007, Duflo et al. 2008, Sauer and Tchale 2009, Fairhurst 2012, Bezu et al. 2014, Franke et al. 2014 and Manda et al. 2015, Droppelman et al 2017), albeit heterogenous, and conditional on farmer wealth and assets (see, among others, Marenya and Barrett 2007, Mugwe et al. 2009, Place et al. 2009, and for a critical review Vanlauwe and Giller 2006).

While studies have found that farmers recognize the value of increased soil fertility (Lambrecht et al. 2015), this understanding does not necessarily imply that farmers also value ISFM technologies. For instance, Ortega at al. (2016) find that farmers significantly discount legume yields in favor of maize yields despite the additional soil fertility benefits provided by legumes. Overall, the adoption of ISFM technologies in Sub-Saharan Africa is still per-

[5] Fertilizer trees improve the condition of soils by fixing atmospheric nitrogen and incorporating it into the soil through their roots and falling leaves. They can also bring nutrients from deep in the soil up to the surface for crops with roots that cannot reach that depth. Fertilizer trees are further useful for preventing fertilizer runoff, and improving water usage for crops. fertilizer trees include Sesbania, Gliricidia, Tephrosia, and Faidherbia albida.

ceived to be low (see, for instance, Wossen et al. 2015, Nkonja et al. 2016 and Nkonja et al. 2017). Sheahan and Barrett (2017), using the World Bank's LSMS surveys for six countries in Sub Saharan Africa, note that while the uptake of inorganic fertilizers and agro-chemicals is not uniformly low (and, in fact, is high in Malawi), there is low correlation between the use of commonly paired inputs (such as fertilizer and hybrid seed; or organic and inorganic fertilizer) at the household, and more importantly, plot-level.

Farmers in Malawi have adopted some ISFM technologies, but few exhibit production characterized by coordinated adoption of multiple practices. In effect, some of the recommended technologies are widely in use, such as intercropping (53 percent of fields) and, as noted earlier, application of mineral fertilizers (55 percent of fields). Other technologies are less used such as organic fertilizers (19 percent of fields), and herbicides/pesticides (2 percent of fields).[6]

CDI aims to increase the adoption of ISFM technologies among smallholder farmers in Malawi, through both extension and improved market access. In this study, we focus on CDI's two primary extension activities: farmer-led demonstration plots and farmer field-days. The annual implementation calendar for these activities follows Malawi's agricultural cycle. In central Malawi, where our study is set, the rainy season starts in November/December and ends in April/May.

In November, CDI sets up demonstration plots in central village locations, close to a primary road, and, according to their own account, on good quality plots. The exact location of the demonstration plot is usually selected through discussion with the local government extension agent and local farmers. Once the location is determined, a CDI extension agent sets up the plot together with a local farmers' club of typically 10 to 20 members. The CDI extension agent continues to guide the farmer's club throughout the season using telephone calls and in-person visits, but the club is in charge of the day-to-day management and the implementation of the various plot activities, such as planting, weeding, fertilizing, and applying pesticides. At harvest, the club members share the proceeds from the demonstration plot.

In March, CDI selects some of the best performing demonstration plots to hold farmer field-days. Farmers in nearby locations are invited to attend these one-day events, with CDI providing transportation to and from the field-day location and a mid-day snack. During the field-day, the CDI extension agent together with the demonstration plot's farmers' club, explain to the visiting farmers how they applied the new technologies. CDI typically holds one field-day per EPA (Extension Planning Areas, a sub-district administrative unit), and

[6] These statistics come from the most recent 2016-17 round of the Integrated Household Survey (IHS, Republic of Malawi, 2017)

invites up to 1000 farmers.

In 2014-15, the growing season under consideration for this study, CDI focused on the following crops: soybean, maize, groundnut and common bean. For these, CDI set up three types of demonstration plots: (1) soybean/maize, the most common type, and; (2) groundnut/maize and (3) common bean/maize. We focus our analysis on soybean and maize, as this is also CDI's primary focus. Almost all farming households in Malawi cultivate maize: In 2016-17, 76 percent of fields were under maize cultivation (IHS, 2016-17). Soybean cultivation has been increasing in central Malawi and is now an important cash crop. In our survey, 40 percent of households cultivated soybean in the 2013-14 season.

CDI defined best practices on the soybean-maize demonstration plots as follows:[7]

- Maize: Use of a high-yielding variety, SC719, optimal plant spacing and seeding practices, regular herbicide (harness and roundup) and fertilizer application (23:21:0+4S and urea), a mulch of crop residues and other transferred biomass, and, on some subplots, the use of fertilizer trees (Tephrosia) and rotation with soybean.

- Soybean: Seed of a high-yielding variety, Squire, treated with inoculants, optimal plant spacing and seeding practices, regular pesticide (cypermethrin), herbicide (harness and roundup) and fungicide (folicur) applications.

All best practice subplots featured ridges, and included the incorporation of organic material through the addition of compost-manure (a mix of decomposed plant residues and livestock manure).

In addition to best practice subplots, each demonstration plot also included control subplots and farmer practice subplots. The latter are understood as "the local method to cultivate a crop"[8], while the former aimed to provide a benchmark for the best practice subplots (using the same variety and planting techniques, but did not include any other external inputs, such as innoculants, fertilizer, pesticides and herbices). Appendix A provides details of the layout.

[7]The best practices for groundnut and common bean are defined as follows: (1) Groundnut: Use of a high-yielding variety, CG7, optimal plant spacing and seeding practices, regular fertilizer (D-compound, single superphosphate and gypsum), pesticide (karate) and herbicide (harness and roundup) applications., and (2) Common bean: Use of a high-yielding variety, Kholophete, optimal plant spacing and seeding practices, regular fertilizer (D-compound), pesticide (cypermethrine) and herbicide (harness and roundup) applications.

[8]In practice, however, these guidelines were not always adhered to, and several farmer practice subplots resembled either control subplots or best practice subplots.

## Sample, randomization and data collected

In 2014, CDI was planning an expansion of their program into two districts in central Malawi: Dowa district, north of Lilongwe, the country's capital, in the central region of Malawi and Kasungu district, North-West of Lilongwe, bordering Zambia. Dowa is a densely populated district with lower than national poverty rate and an average climate. Kasungu has a lower population density, a large hinterland and higher than average poverty rate.

Together with CDI, we selected two EPAs (Extension Planning Areas, a sub-district administrative unit): Chibvala in Dowa district and Mtumthama in Kasungu district. The 2014 village census listing of the District Agricultural Offices included 360 villages in these two EPAs. We randomly selected 250 from the 303 villages which counted at least 50 households, stratified by EPA.[9] Half of these villages, again randomly selected and stratified by EPA, were assigned to the treatment group and the other half to the control group. The villages in the treatment group were invited to form farmer clubs and to participate in CDI's program. We provide evidence of the success of this randomization process for our study sample in Appendix Table 1.[10]

Farmers formed clubs in 91 out of 125 treatment villages (in our study sample, this will be 47 out of 55). Observing this club formation process it appears that villages who formed clubs tend to be further away from markets and have a history of community action (we will confirm this observation, and discuss its implication, in the next section).

Seventeen of the 91 villages received farmer-led demonstration plots during the 2014-15 growing season. These 17 villages were selected by CDI. By CDI's own account, these were villages with some familiarity with agricultural extension services, located in an accessible location and where people were in "unity". All but five villages received a soybean/maize demonstration plot, the remaining five received either a groundnut/maize plot or a common bean/maize plot. Appendix table 2 compares the demonstration plot villages with the other treatment villages, and reports few significant observable differences between these two sets.[11]

[9] As CDI works through farmer clubs and the functioning of these clubs requires a minimum village size, we excluded the villages with less than 50 households.

[10] Appendix Table 1 compares village-level descriptive statistics for the control villages (Columns (4) through (6)) and the treatment villages (Columns (7) through (9)). Column (10) presents the P-values of a t-test with unequal variances. We note three areas of differences between the two groups: ethnicity, size of civil organisations, and daily wage rates. Note that the sample includes only the villages who were revisited one year after baseline (see the data collected section below).

[11] Appendix Table 2 compares the demonstration plot villages with the other treatment villages. Recognizing the small sample sizes, we note that the demonstration plot villages do not appear to very different from the non-demonstration plot villages, although the former appear to be somewhat more accessible. Note that the sample includes only the 55 treatment villages who were revisited one year after baseline, including

In March 2015, CDI invited all farmer clubs in the 91 villages to a farmer field-day. Two field-days were held in the study area. Farmers in Mtumthama EPA were invited to a local farmer field-day at the best performing CDI farmer-led demonstration plots in the EPA. Farmers in Chibvala EPA were invited to join a farmer field-day in a neighboring EPA (Lisasadzi EPA), due to, according to CDI's account, the lack of exemplary demonstration plots in Chibvala EPA itself. Both field-days took place at a soybean/maize demonstration plot.

## Data collected

We collected data at baseline, before the treatment villages participated in the program activities (in Fall 2014), and one year later (in Fall 2015). The baseline was conducted in all 250 villages in the sample, while the data collection the following year included 100 villages.[12]

Before collecting baseline data, we generated a census of all households in the 250 villages as well as a census of all CDI club members in the treatment villages. We used these two census lists to draw a sample of 10 households for each village: In the control villages and the treatment villages without a club, we randomly selected 10 households from the village census. In the treatment villages with CDI clubs, we stratified the sample and sampled five households not participating in a CDI club and five participant households.[13] One of the five households sampled was the household of the lead farmer of the club, whom serves as the point of contact between CDI and the club. The other four CDI households were randomly selected from the list of households who belong to the CDI club. We discuss self-selection into clubs in the next Section. Club members are wealthier, more educated and better connected than non-club members.

At baseline, we conducted a village survey, a household survey and collected and analyzed soil samples from farmers' plots and demonstration plots. One year later, we followed up with household surveys; creating a household panel dataset.[14] Between these two rounds

the 10 additionally selected demonstration plot villages (see the data collected section below).

[12] These 100 villages were selected as follows: First, we selected selected 90 villages randomly from the 250 sample villages, stratified by EPA and treatment status. These 90 villages included 7 villages with demonstration plots. Then, we included an additional 10 villages which had been selected as demonstration plot villages by CDI (as to include all 17 villages which were selected as a demonstration plot locations).

[13] In case of multiple CDI clubs in a village, we selected the club to be included in the study randomly. In terms of the treatment, all CDI clubs are invited to the farmer field day while only one club was engaged per demonstration plot (this would be the same club which we interviewed).

[14] The attrition rate is 5% - specifically, there were 51 households who were present at the baseline who were not present in the follow up survey. The households who left the sample are uniformly distributed

of data collection, we collected agronomic data at the demonstration plot sites on a weekly basis. We also conducted a series of focus group interviews and interviewed extension agents. We discuss these data sources below.

### Village survey

We administered a village questionnaire at baseline in each of the 250 villages with a knowledgeable individual, often the village head or the secretary to the village head. This village questionnaire covered information on the village's distance to paved roads, national highways, (seasonal) markets and other services (such as banks). In addition, we collected demographic information (number of residents, ethnic distribution), and information on access to government and NGO extension, civic organizations and the price of casual agricultural labor in different seasons. We noted the location of the village center using GPS.

### Household survey

We conducted a household survey among 2500 households in 250 villages at baseline, and among a subset of 1000 households in 100 villages one year later. The survey was collected in the months of October and November, about five months after harvest and right before planting for the next season. We interviewed the head of the household.

At baseline, we collected data on household composition, groups, networks and information sources, landholding, marketing, subsidies and credit[15], and assets. At both baseline and one year later, we collected information on the adoption of ISFM technologies and yield expectations. One year after baseline, we also collected data on knowledge of ISFM technologies. Given the focus of this study, we detail the latter three modules below.

**Adoption of ISFM technologies** At baseline, we collected information on current use of ISFM technologies using an input-output plot-level questionnaire (pertaining to the previous, 2013-14 season). We focus on the technologies introduced by CDI and include

---

geographically and in terms of treatment status. The households who left the sample have household heads who are slightly younger (0.01 years – significant at the 5% level) and slightly more educated (0.05 years – significant at the 10% level) but do not differ in terms of household composition and asset wealth. To keep the sample size intact, these 51 households were replaced in the follow up survey using the random sampling methods outlined above.

[15] We asked the household head to list all lines of credits taken up in 2013-14, the terms of the credit, and where he/she would go if he would like to obtain credit for the next (2014-15) season (and how likely he/she believes credit can be obtained from this source).

seed treatment, seed selection, plot lay-out (intercropping, rotation, fallow, etc.), fertilizers (inorganic, organic and fertilizer trees), and other inputs: pesticide, herbicide, fungicide. To obtain a longer term picture, we also collected information on ISFM technologies used in the past five years, again, with a focus on CDI technologies, asking the farmer whether in the last five years, they had (ever) used a particular technology.

One year later, we repeated the input-output questionnaire (this time focusing on the 2014-15 season), and added a new module on adoption plans. In the latter, again, we focused on CDI technologies and asked whether or not the respondent plans to adopt a particular technology (in the 2015-16 season), and if so, followed up with details, and if not, asked for reasons for non-adoption. Recall that the household survey was conducted in the month preceding planting. Hence, most respondents were comfortable with these forward-looking questions and had no difficulty in responding.

**Knowledge about ISFM technologies** This module was only included in the follow-up survey, one year after the baseline (and not at baseline). We build on Kondylis et al. (2015) and incorporated twenty questions designed to assess knowledge about the ISFM techniques introduced by CDI. Directly eliciting knowledge, even though perhaps not often included in standard surveys, has become more and more common.[16]

The questions we, together with CDI, created covered ISFM practices for soybean, groundnut and maize. All questions had a correct response. Responses were true/false, multiple choice or a numerical. Questions ranged from listing the general benefits of certain ISFM practices, such as the benefits of growing soy bean in crop rotation, and covering the soil with crop residues, as well as knowledge about how-to-apply ISFM practices including: how many weeks after planting should you apply urea fertilizer on maize; what chemical is best for controlling soy rust; where on the field should one plant fertilizer trees; and when mixing inoculant, how many table spoons of sugar should one add to the inoculant bag. The questions are listed, in the format presented to the farmers, in the Appendix. We code the answers as correct/incorrect and compute a total knowledge score (out of 20). Note that questions were designed to reflect topics that are important parts of the promoted technologies. [17]

[16] Screening papers presented at the latest AAEA conference in Chicago (in 2017) reveals that there were around 20 papers which used some form of knowledge elicitation, among others, Cai et. al. (2017) and Pan et al. (2017). See also Laajaj and Macours (2017A) for a critical discussion of various measures of ability, skill and knowledge.

[17] Respondents could also opt for a "don't know" response, which we coded as incorrect.

**Yield expectations** We build on Delavande et al. (2010 and 2011), Dillon (2014) and Maertens (2017) to elicit yield expectations at both base and endline. We focus on soybean, groundnut and maize. At baseline, we asked the respondent: "Imagine that you would cultivate maize this coming year (and that maize is the only crop on the field, i.e., no inter-cropping), how much maize do you think you would harvest on one acre of land" We recorded the answer in 50 kg bags of shelled or unshelled maize. We then repeated these questions for soybean (in 50 kg bags of shelled soybean) and groundnut (in 50 kg bags of unshelled, dried groundnut).[18][19]

### Focus group interviews and interviews of extension agents

We conducted focus group discussions in ten villages before the CDI program, and one and two years after the program. We interviewed CDI clubs whom had been invited to a field-day and clubs who managed demonstration plots. We followed best practices (see Morgan 1996 and Krueger and Casey 2008) and focused on learning about agricultural technologies, the club's activities and challenges faced, and relationships with extension agents.

We conducted semi-structured interviews with two government extension agents and two CDI extension agents in our study area, before the program and one and two years after the program. We focused on the constraints and opportunities facing extension agents, and their relationship with the farmers.

[18] In our baseline data, while mono-cropping was the norm (with 75% of the plots mono-cropped), inter-cropping is common. Due to the complexity in generating per-acre beliefs on inter-cropped fields, we asked the respondent to imagine a mono-cropped field. The unit was determined in qualitative interviews preceding the data collection as most common unit people think about for the crop. In addition, we recognise the difficulty in imagining the exact size of one acre of land, and in the formulation of this question we often referred to a 70 by 70 feet area or provided a comparison field in the village. However, we do expect measurement error due to the lack of ability to imagine exactly the size of one acre, and also asked the respondent for the expected yields on a particular field, instead of a per-acre basis (see also Bevis and Barrett 2016).

[19] At endline, we expanded this module. To obtain a probability distribution, we first asked the respondent to describe the best growing conditions, average conditions and worst conditions he/she could imagine for maize. In this round, we distinguished between hybrid maize and local maize. Respondents, in response, often noted variance in weather, pest pressure etc. Then, we asked him/her to state how much maize he/she would harvest under the best condition, the average condition and the worst condition, respectively. Finally, we asked the respondent to distribute ten equal size stones (each representing a 10 percent probability) in three equal-sized circles drawn on the ground, the first circle representing the best condition, the second the average conditions and the third the worst conditions. We repeated these questions for soybean and groundnut.

### Field observations on demonstration plots

We visited the demonstration plots two weeks after planting to record germination and record activities and inputs used to up to that date. Data on agronomic practices were recorded via a phone call with the lead farmer on a weekly basis between planting and harvesting. During this weekly phone call we recorded any activity that had taken place, such as applying fertilizer or other inputs, and the number of club members and other visitors present for the activity (including whether the CDI extension agent was present). Rainfall gauges were mounted on each demonstration plot and the lead farmer was trained to record rainfall on a daily basis.

At harvest, we visited the demonstration plots and collected crop yield data. We recorded the stand count at harvest, the total biomass, grain yield and stover or leafy biomass. Grain moisture content was determined using a Mini GAC plus moisture meter. It is important to note that the club members were present during these on-field activities, and hence, are expected to have good idea of the planting and harvesting counts.

### Soil sampling and analysis

The key indicators of land fertility in the study area are soil pH and organic matter content (see Snapp 1998). We collected soil samples from a total of 225 farmers' fields in addition to the 19 demonstration plots during November-December 2014 and 2015.[20] For each field, we first recorded the cropping history and then asked the farmer to guide us to the field. After recording the GPS coordinates of the field in a central location and walking around the field to record the field area, we collected two soil samples at 0-20 cm soil depth. These samples were then mixed to make a composite sample. After collection, the soil samples are put in soil sampling bags and taken to the Bunda College Soil and Plant Analysis Laboratory for analysis. If the soils were wet upon arrival at the laboratory, the samples were first air dried. When dry, we sieved them through a 2mm sieve and recorded the soil texture using the hand feel method.

[20] The farmers were selected as follows. First, we selected all ten sample farmers who live in villages where a CDI demonstration plot was set up. Second, we randomly selected 20 treatment villages and 9 control villages, and approached all 10 farmers in each village for soil sampling. Third, we selected 10 villages purposefully, for their relatively higher share of female-headed households and collected a soil sample from all ten sample households in these villages. This results in a total of 560 farmers, of which 225 live in villages which were covered by our follow up survey. As many farmers cultivate more than oen field, we asked farmers to identify the field they would be most likely to try new technologies on. Farmers are more likely to select fields they own, fields of mixed soil texture, and fields with a higher incidence of soil erosion and nutrient depletion (Regression results are available on request).

We use the SoilDoc program to analyze the sample pH, nitrate nitrogen ($NO_3^-$), inorganic phosphorus (P), sulfur (S), exchangeable potassium (K) and electrical conductivity (EC), and active carbon (C). See Gatere at al. (2013) and Weil and Gatere (2015) for an introduction to SoilDoc. Note that we did not measure the total organic carbon matter, a measure of carbon contained within the soil organic matter, and generally accepted to be a good summary measure of overall soil fertility. Instead, we measured active carbon, which compared to total organic carbon, is more sensitive to management effects, and more closely related to soil productivity and biologically mediated soil properties, such as respiration, microbial biomass and aggregation (Weil et al. 2003).

## Descriptive statistics

Appendix Table 1 introduces the 100 villages which were re-visited one year after the baseline (in Columns (1) through (3)) - the statistics reported refer to the baseline year. We note that the villages exhibit considerable variation in distance to services, with some villages situated almost as far as 40 km from local markets. In terms of composition, most villages are small, just 62 households, on average, and most have only one ethnic group. The villages are visited by government extension agents, on average, eight times per year, and by non-governmental extension agents, on average, 16 times per year. The variation in the number of yearly visits is substantial, with some villages benefiting from weekly visits, while others reporting no visits. Most villages have one or two civic organizations active.

The households in our study area are land-poor and dependent on rainfall agriculture. Table 1 - Panel A - introduces the households. We present baseline statistics of the households in the 100 villages who were re-visited one year after baseline.

The average household head is 42 years and has 4.5 years of education. About 18 percent of household heads are female, and the average household has 5.22 household members. Credit constraints are likely important for these farmers. Only 17 percent of the respondent stated that they had taken out credit the previous season. Respondents report receiving information from both government extension agents and fellow farmers. About 40 percent of respondents report having interacted with government extension agents, while another 20 percent report to interact often to very often. For 30 percent of the respondents, extension agents are one of the main sources of information. This is consistent with the accounts received during the focus group interviews when farmers mentioned to learn about new agricultural technologies through own experimentation, from other farmers and from extension agents.

On average, households own 3.5 acres of land. While this figure excludes outliers above the 95 percentile, it might still appear high. It is important to note however that the median field size is small, 1.5 acres, and likely to be an over-estimate (these are self-reported acreages, which are often over-estimated in the case of smaller plots, see Bevis and Barrett 2016 for a discussion).

Plots of land are small and population density is high in the area, according to the respondents' own account often the result of generations of plots being sub-divided for inheritance. A lack of land can further reduce land quality, as leaving fields fallow, or using crop rotation, might no longer be options for many households. Soil fertility in the study area is low, and declining. Soils are classified as Ferralsols, Lixisols and Plinthosols (FAO Harmonized World Soil Database).[21] A common feature of these soil types is that they depend on the addition of organic and inorganic matter to improve soil structure and overall fertility. Lacking these, soil fertility will be low. Our respondents report soil fertility problems (see Table 2 - Panel B): 80 percent of farmers perceive the average soil fertility to be stagnant or declining. Common reported problems are: soil erosion (by 47 percent of households), water logging (by 23 percent of households) and nutrient depletion (by 57 percent of households). Acidity and salinity is not a main concern.

The soil sample analysis results, summarized in Table 1 - Panel C - confirm these perceptions. Soils have an average Ph of 6.12 (with 70 percent of soils tested within the optimal range, between 5.5 and 6.5), so only slightly acidic. Active carbon is, on average, 423 mg/kg soil, and is low to very low in 30 percent of soils tested (with a concentration of less than 350 mg/kg soil), and at a medium level in another 40 percent (with a concentration between 350 and 500 mg/kg soil). This would indicate that organic carbon is mostly sufficient to maintain soil structure, but still low. These findings are consistent with widespread nutrient deficiencies. All soils tested are Nitrogen (N) deficient (less than 42 mg $NO_3^-$ /kg soil), 77 percent are Sulphur deficient (with concentration is less than 10 mg/kg soil), 52 percent are Phosphorus (P) deficient (with concentration is less than 0.3 mg/kg soil) and 33 percent are Potassium (K) deficient (with concentration less than 20 mg/kg soil). All-in-all, over 50 percent of soils were deficient in three or more nutrients. The intra-class correlation between observations of the same village is below 0.5 for measures of Nitrogen and Potassium, indicating significant within village variation in these measures.

Table 1 - Panel D reports farmer cultivation practices in 2013-14. Almost all respondents cultivated maize (of which 62 percent opted for a hybrid variety) and less than half have

[21] The former are old, weathered soils, sandy-loam free-draining with resulted low nutrient content and possibly acid Ph. Lixisols are more sandy-textured version of Ferralsols and hence more subject to erosion. Plinthosols typically have a hardened layer of iron and/or alumninum deposits impeding water flow and root development.

cultivated soybean, common bean and groundnut. Over 80 percent had used crop rotation (in the last five years) while about half reported using intercropping in 2014-15. A large number, 88 percent, had used mineral fertilizer (Malawi has a large-scale mineral fertilizer subsidy program targeting small-holder farmers). Animal manure had been used by 30 percent (but compost was not common), 14 percent had incorporated crop residue in the soil, and 9 percent had planted fertilizer trees. The use of other inputs is not very common. Less than 5 percent had used pesticides, herbicide or fungicide; and only two soy growing farmers had inoculated the seed.

Finally, Table 1- Panel E reports on the farmers' baseline yield expectations for maize and soybean. Farmers expect to harvest, on average, 3,480 kg/ha (or 29 kg bags per acre) of shelled maize. This is significantly larger than the average yield on the farmers' plots in 2013-14, which was 1,750 kg/ha for mono-cropped plots. However, it is important to keep in mind that: (i) the yield expectation distribution is not normal, with a long left tail - in effect the median of the distribution is 3,088 kg/ha, and (ii) the beliefs are reflect also perceived acreage, which, in our data, are over-estimated for smaller plots and under-estimated for larger plots (we know this as for a sub-set of the plots, we also have the GPS measured acreage). Farmers expect to harvest, on average, 1,608 kg/ha or (13 kg bags per acre) of shelled soybean. This is larger than the average actual yield in 2013-14 (312.5 kg/ha, on mono-cropped plots). Again, the same disclaimers apply and it should be noted that the median yield expectation is 1,235 kg/ha.

As a comparison, Table 2 presents the yields obtained on the demonstration plots in 2014-15. We focus on the maize and soybean subplots. Maize grain yield was variable and ranged from 452 kg/ha under control treatment to 8,990 kg/ha under best practice agronomy. The latter is within the range of potential yields for maize, ranging from 6,000 to 14,000 kg/ha (depending on the variety, see MAIFS, 2012). Overall, the use of best practice agronomy practices increased maize yield by 62 percent and 25 percent over the control and farmer practice treatments respectively.[22] However, no differences were observed between best practice agronomy treatments with and without fertilizer trees.

Differences in grain yield of soybean between the treatments and sites are also significant. Yields range from 0 (crop failure) to 2,218 kg/ha. The use of best agronomy practices increased the yield of soybean by 50.4 percent over the control.[23] Overall, the yields of soybean are somewhat lower than the potential yield of 2000-4000 kg/ha, but in the range of the attainable yields on smallholder farms (1,500-2,500 kg/ha) with good agronomic practices (MAIFS, 2012).

[22] Based on within demonstration plot analysis, excluding outliers.

[23] Based on within demonstration plot analysis, excluding outliers.

The lower than attainable yield performance on some of the demonstration plots could, among others, be attributed to the poor rainfall distribution in the 2014-2015 cropping season. The average total rain on the demonstration plots in 2014-15 was 665 mm (st. dev. 196 mm). In addition, it should be noted that not all demonstration plots followed the Best Practice Guidelines to the letter.

To summarize, respondents were well aware of the low soil fertility in the region, and note significant nutrient depletion in their own soils. Accordingly, yield expectations are low, and well below both attainable yields (albeit considerably above the actual yields of the previous season, 2013-14). Despite this awareness, adoption of ISFM technologies has been relatively low, especially of certain information intensive technologies such as the use of pesticides, herbicides, fungicides, fertilizer trees, compost, and inoculation. Our central hypothesis is that this low uptake is partially due to a lack of knowledge, and that exposure to demonstration plots and field-days can remedy this lack of knowledge.

Note that a demonstration plot effectively combines three sources of learning (self, social and from external agents), and during the baseline focus group interviews farmers in CDI clubs, by their own account, noted to be keen on establishing a demonstration plot in their village to learn about production processes. Demonstration plots have been central to much of Malawi's extension history (see Knorr et al. 2007), and farmers, prior to the program, stressed that the best way to learn is to work on a demonstration plot together.

When we spoke with the farmers (during the focus groups) one year, and even two years after they had begun working on a demonstration plot together, many were able to recall the exact names of the ISFM inputs used, report the amounts used and explain how the inputs should be applied. They stated feeling "comfortable" with the techniques. Farmers who had only attended the field-days stated that they were impressed by the productivity of the crops presented at the field-day, and reported that they learned about the importance of herbicide and pesticide (and in some cases inoculants) for soybean, as well as plant spacing and the importance of using crop residuals for mulching and plant spacing. In contrast, few field-day participants were able to recall the details: which input was used; the amounts required; and the method of application. This is consistent with what farmers told us before the CDI program started in the region: while learning from extensions services at field-days and the radio is common, they also noted that they rarely immediately adopted the new technologies after visiting a field-day or hearing something on the radio, as through these channels they are less aware of the intricacies of how the technologies work and less certain as to how they can be applied on their fields.

We tested the knowledge of farmers after the treatment group had participated in the program. In Table 3 we present summary statistics for this knowledge test (recall, the test was administered one year after baseline). The overall knowledge score is 7.87 - out of 20 (with a standard deviation of 2.30). We note that while most respondents are aware of the general benefits of soy, fewer know the details of the production process in terms of which pesticides and fungicides one should apply following best practices. The share of correct answers drops even further - to under 10 percent - when we ask the respondent to tell us about the details of soybean input preparation and application. For maize, a crop with which farmers have extensive experience, farmers seem to be aware of certain ISFM technologies, such as the use of crop residues and fertilizer trees but have limited knowledge of the details of the production process as well.

In the next section, we will test for the impacts of the demonstration plots and field-days on knowledge and adoption plans.

# Impact of the program on adoption and knowledge

To estimate the impact of the CDI program, one would ideally run a regression such as specification (1) linking outcomes $Y_{ij}$ of farmer $i$ from village $j$, on whether or not the farmer participated in demonstration plots $D_{ij}$ or field-days $F_{ij}$:

$$Y_{ij} = \alpha_0 + \alpha_F D_{ij} + \alpha_D F_{ij} + \epsilon_{ij} \tag{1}$$

However, while being invited to a farmer fieldday is randomized at the village level, a critical aspect of participation is a choice: Farmers have to sign up for CDI clubs in order to become eligible for the CDI activities. As one is unlikely to be able to control for all relevant confounding factors - many are unobservable to the researcher such as, climatic factors and personal attributes - we might expect a correlation between $\epsilon_{ij}$ and $F_{ij}$ and $D_{ij}$, resulting in omitted variable bias.

Appendix Tables 2 and 4 shed light on this participation decision. Recall that 47 out of 55 treatment villages formed clubs. Appendix Table 2 presents, in Columns (8) through (14), the descriptive statistics of villages which formed CDI clubs and villages which did not form CDI clubs, along with the results of the t-test. While the sample sizes are small, the CDI villages appear to be different from the eight villages that did not form clubs: the latter are better connected, a smaller acreage under soy, and have fewer existing organizations and groups.

In Appendix Table 3, Column (11), we show the result of t-tests testing the baseline differences between club and non-club members in the villages which formed clubs. Club-members are different from non-club members in many dimensions, and that this difference is both statistically and economically significant. Compared to non-club members, club-members are better educated, have larger families and more land, and are also more likely to take agricultural credit.[24][25]

Our estimation strategy takes this self-selection into account by constructing two comparable samples: the sample of households which received the CDI program and a comparable sample of households, in the control villages, which do not received the program. Using this approach, we essentially assume that the treatment villages are comparable to the control villages; and we can use the latter to construct a similar sample (we presented evidence of the similarity between treatment and control villages in Appendix Table 1).

To do this, we first predict the probability of club membership using the treatment villages sample only:

$$P(club_{ij}) = \beta_0 + \beta_1 X_{ij} + \beta_2 X_j + \mu_{ij} \tag{2}$$

The control variables used in this first step include all village level characteristics included in Appendix Table 1 ($X_j$), all household level characteristics included in Appendix Table 3 (and the square terms of the non-binary variables) and baseline adoption indicators ($X_{ij}$). We do a reasonable job predicting club membership - with about 70% correctly classified.

We use the estimated coefficients to then do an out-of-sample prediction for the control villages. We summarize the results of this process in Appendix Figure 1 where we present the resulting kernel distribution of this predicted probability for both the treatment villages and the control villages. As expected, the mass of the distribution of the control group is situated to the left of the median of the treatment group: This is due to our sampling design - we stratified the sample in the treatment villages and likely over-sampled club members.

[24] This is consistent with Laajaj and Macours, (2017B) who find that, in Siaya, Kenya, the farmers that villages select to manage demonstration plots are notably different from randomly selected farmers: the village selected farmers were higher educated, with higher skill levels, and had more assets and larger families. They were also more likely to already having adopted more of the recommended ISFM practices, and put in more effort into the demonstration plot activities.

[25] Appendix Table 3 presents the same statistics, but comparing households in treatment and control villages. We find a few statistically significant differences between households in the control villages and in the treatment villages: households in the treatment villages own more land, on average, and were more likely to indicate government extension agents as one of the main sources of information. However, recall that the sample was constructed differently in the treatment villages. While in the control villages, a random set of 10 individuals were selected; in the treatment villages with clubs, the sample was stratified and both club members and non-club members were sampled. If these two sets of individuals are different, and the two groups were drawn from different-sized populations, we would not expect this balance test to show equality.

We then create two groups in each set of villages: (would be) club members and (would be) non-club members using the cut-off value of the predicted probability of 0.5. This includes 241 individuals in the treatment group and 173 individuals in the control group. Appendix Figure 2 presents the distribution of the predicted probability of these individuals. This time, the distributions are comparable. A two-sample Kolmogorov-Smirnov test for equality of distribution functions cannot reject equality between them with a P-value of 0.22.

We now compare the club members in the treatment villages with the would-be club members in the control villages using:

$$y_{ij} = \alpha_0 + \alpha_1 T_{1ij} + \alpha_2 T_{2ij} + \varepsilon_{ij} \tag{3}$$

Where $y_{ij}$ is the outcome variable of farmer $i$ from village $j$; $T_{1ij} = 1$ if the individual is in a club which managed a demonstration plot, and $= 0$ otherwise; $T_{2ij} = 1$ if the individual is in a club which was invited to a field-day, and $= 0$ otherwise. We clustered the errors at the village level, and used a bootstrap procedure to account for the two stages in the estimation process (see Abadie and Imbens, 2009).[26]

The dependent variables $y_{ij}$ include the planned adoption (2015-16 season) of: soybean, inoculation of soybean, groundnut, hybrid maize, herbicide, pesticide, fungicide, inorganic fertilizer, fertilizer tree, intercropping, animal manure, crop residue and compost and whether or not each one of the questions in the knowledge test was answered correctly. We also compute an adoption score (out of 13) and knowledge score (out of 20).

Selection into $T_{1ij}$ (as opposed to $T_{2ij}$) might still be a concern. However, as noted earlier, the demonstration plot villages are comparable to the other treatment villages (see Appendix Table 2). Appendix Table 4 shows that this lack of selection also holds at the individual level: demonstration plot farmers are comparable to the CDI club members who did not manage demonstration plots.

Tables 4 and 5 present the results. Table 4 shows that demonstration plot participation increases adoption of ISFM practices by 0.79 points, which is about 22 percent (Column (1)), while being invited to a field-day does not produce such (statistically significant) result. Columns (2) through (14) present the results of a series of linear probability models. We note that being a member of a demonstration plot club increases the chances of inoculating

[26] This method resembles propensity score matching, but does not, as in propensity score matching, directly compare individuals in the treatment group with individuals in the control group with similar propensity scores. This method is not (very) feasible in our context: most of the households in the treatment group have a very high predicted probability of being including, making the overlapping support assumption which underlies propensity score matching problematic. Furthermore, it was unclear to us what an average effect, over the various predicted probabilities, could be interpreted for policy purposes.

soybean (at the 5 percent level), using hybrid maize (at the 1 percent level) and planting fertilizer trees (at the 5 percent level).

Table 5 present the impacts on knowledge. Column (1) presents the effects on the knowledge score out of 20 - with one point for every question the farmer answered correctly. Columns (2) through (21) present the result of 20 linear probability models, taking as a dependent variable whether or not the farmer answered each of the 20 questions correctly.

Participating in a demonstration plot improves knowledge. Table 5 reports an increase by 0.63 points, which is about 8 percent (P-value of 0.11), with statistically significant increases in knowledge of inoculation and pesticides in particular (note that we do spot a negative effect on question (14) - but the negative, tricky, formulation of that question cautions against over-interpretation). In contrast, being invited to a field-day does not (statistically significantly) alter one's knowledge.

The results so far seem to suggest that the field-days were not very effective. Results of focus group discussions however indicate that field-day participants learned about the production processes of some of the more labor-intensive ISFM technologies, such as mulching and optimal plant spacing. Farmers reported that mulching, for instance, was a useful technology to combat striga (a common weed in maize fields which can cause heavy crop losses), and also the very common drought spells. This suggests that these farmers, who are likely constrained as to what they can focus on during one day, focus on these technologies that they are most likely to successfully implement. Many of the recommended inputs – pesticides, herbicides and inoculants – are not available in local markets. As few households own cars or motorbikes and public transport is limited, distance could represent an additional barrier to participation in input markets. Many farmers also noted that, even if such inputs were readily available, they would not have the funds available to purchase these inputs. Hence, field-day participants, by their own accounts, often focused on the labor-intensive technologies, rather than the credit-intensive technologies. This would suggest that both groups are learning, but learning should be considered a choice, a choice which is constrained by factors such as credit and time.

In the next section, we present a learning model which follows this narrative, capturing the motivation and constraints to learn about ISFM technologies.

As a robustness check, Appendix Tables 5 and 6 present the results including the first-stage set of control variables in the second stage equation (3) and Appendix Table 7 presents the results of a farmer fixed effects estimation. The results of these specifications are broadly consistent with the ones presented in Table 4 and 5.

The validity of our estimation depends on our ability to create a comparable control group of would-be club members in the control villages. Systematic errors in our first stage could result in over-classification (placing people in the comparable group whom would not have joined, possible biasing our estimates upwards), or under-classification (placing systematically better people in the comparable group, possibly biasing our estimates downwards). However, the fact that the predicted distributions of treatment and control group largely overlap (Appendix Figure 2), and that adding additional control variables does not oppose our results is encouraging.[27]

We conclude this section with a note on participation in the program. Tables 4 and 5 present the intent-to-treat effects. In the case of demonstration plots, this does not matter much: all CDI club farmers who were invited to participate in demonstration plot activities actually did. In the case of field-days, the estimated effect should be interpreted as the intent-to-treat effect, as not all clubs invited to the farmer field-days participated.

In Appendix Table 4 we presents the P-values of a t-test, comparing farmers who attended field-days, which those CDI club members who did not attend field-days (Column (7)). We note considerable differences: Farmers who attend field-days are higher educated, own more land and are overall wealthier and better connected, and are more likely to use credit. This is consistent with the observation that in most clubs, it is the lead farmer who attended the field-days. Our estimates should be interpreted as the average effects, across all club members, immediately after the intervention. As CDI requested the club members who attended the field-day to share the information with the other members, this, in a way, is also a measure of the success of this process. As we discuss in the concluding section, in future research, which we aim to conduct five years after the first intervention, we aim to provide a more nuanced and complete picture of this social learning component.

# A model of learning

We build on Foster and Rosenzweig (1995), Hanna et al. (2014) and Nourani (2017), to model the farmer's learning and adoption decision as an optimal portfolio choice with multiple objects of learning under a range of initial endowments. Since we do not allow for borrowing, variation in the initial endowments captures the degree of a farmer's credit constraints.[28]

**Yields** We introduce three production technologies: a capital-intensive technology in-

[27] Appendix Tables 5 and 6 are demanding specifications; and many of the bootstrap rounds did not converge (see also Horowitz 2001).

[28] We abstract away from the notion that knowledge decays over time, which would empirically result in a false acceptance of the null hypotheses for in particular demonstration plot farmers who might have learned earlier. We also do not model the fact that demonstration plot participants might share the proceeds of the demonstration plot, which might further incentive them to learn.

dexed $K$, a labor-intensive technology indexed $L$ and a traditional technology which represents a risk-free technology. Each risky technology $K$ and $L$ has average per-acre payoffs (yields) of $\mu_j$ ($j \in \{K, L\}$) — yields associated with the risk-free technology are normalized to one. Furthermore, we assume that the yields from the capital-intensive technology are higher than the labor-intensive technology: $\mu_K > \mu_L > 1$.

**Two-stage approach** The farmer first establishes a belief of the yield (we assume that the farmer does not need to learn about prices), and then invests resources to generate knowledge, i.e., learn about the production process, and makes decisions as to which technologies to adopt.

**Learning about yields** We assume that the true value of $\mu_j$ is unknown to the farmer. Let the prior belief about $\mu_j$, $\hat{\mu}_j$, be normally distributed, centered around the true value, with variance $\sigma_\mu^2$. So each farmer's prior represents a draw of this distribution.

$$\hat{\mu}_j \sim N(\mu_j, \sigma_\mu^2) \tag{4}$$

When observing yields on the demonstration plot, either in the village, or at a field-day, the farmer receive an unbiased information signal $\upsilon_j$. This signal is the sum of the true yield $\mu_j$ plus a normally distributed noise term:

$$\upsilon_j = \mu_j + \eta_j \quad \text{with} \quad \eta_j \sim N(0, \sigma_\eta^2) \tag{5}$$

The variance in the noise term, $\sigma_\eta^2$, can be farmer-dependent, and significant, if the farmer believes that the soil and climatic conditions of the demonstration plot are dissimilar to his own conditions. However, to maintain simplicity, we will abstract from this farmer-dependency in our notation. Assuming the farmer uses Bayesian updating, then noisier signals will be down-weighted in posterior beliefs. Posterior beliefs, $\mu_j^p$ are characterized by:

$$\mu_j^p = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\mu^2}\hat{\mu}_j + \frac{\sigma_\mu^2}{\sigma_\eta^2 + \sigma_\mu^2}\upsilon_j \tag{6}$$

Thus, the posterior beliefs will decrease (relative to the prior) if the signal received during the field day is less than the prior and increase otherwise. However, the degree of change in posterior beliefs depends on the size of $\upsilon_j$ and the farmer's perception of the relative noisiness of the field-day signal.

Equation (6) implies that the posterior belief represents the weighted average between the prior belief and the signal received. In regression terms, this would imply that for farmer $i$ the posterior beliefs of technology $j$ can be expressed as:

$$\mu_{ij}^p = \alpha + \beta_1\hat{\mu}_{ij} + \beta_2\upsilon_{ij} + \varepsilon_{ij} \tag{7}$$

**Production process** Production requires the use of inputs (e.g., amounts and timing of fertilizer, herbicide, labor). If inputs are inaccurately applied farmers incur a *knowledge penalty.* Specifically, let $\theta_j^*$ indicate the optimal amount of input required for technology $j$. If the farmer applies input $\theta_j$ instead, he incurs a (per-unit) loss equal to $(\theta_j - \theta_j^*)^2 < 0$ for all $\theta_j \neq \theta_j^*$.

**Learning about the production process** The optimal input use , $\theta_j^*$, associated with technology $j$ is unknown. Let the belief, $\hat{\theta}_j$, again be normally distributed, centered around the true value ($\theta_j^*$):

$$\hat{\theta}_j \sim N(\theta_j^*, \sigma^2_{\theta_j}(e_j)) \mid e_j \in \{0,1\}, \tag{8}$$

Note the dependency of the belief on learning effort, denoted $e$. In particular, the beliefs are more precise if the farmer applies a discrete learning effort, $e_j = 1$ (compared to the situation where the farmer applies no learning effort, i.e., $e_j = 0$):

$$\sigma^2_{\theta_j}(e_j = 1) < \sigma^2_{\theta_j}(e_j = 0) \tag{9}$$

Note that when the farmer's belief over the target input is imprecise, the knowledge penalty will be large in expectation ($E[(\theta_j - \theta_j^*)^2] = \sigma^2_{\theta_j}(e_j)$) — and as the farmer gains production knowledge, his knowledge penalty decreases in expectation.

**Payoffs** The farmer holds initial wealth $w_0$ and is tasked with choosing the optimal amount of wealth to invest in each production technology, $x_j$, each unit of which costs $p_j$ to purchase. At harvest, the farmer receives the following payoff:

$$w_0 + x_1(1 - p_1) + \sum_{j \in \{K,L\}} [x_j(\mu_j^p - p_j) - \left[x_j(\theta_j - \theta_j^*)\right]^2 - \overline{e}_j 1(e_j = 1), \tag{10}$$

where $\overline{e}_j$ represents the cognitive cost associated with gaining knowledge of production technique $j$ and only contributes to payoffs when learning effort is applied (1 represents the indicator function). The amount invested in the traditional technology is denoted $x_1$, and its cost $p_1$ (recall that the yield of the traditional technology was normalized to 1).

**Expected payoff maximization** Given the presence of credit market imperfections, the farmer's problem will be one of maximization of *expected payoffs* given a budget constraint. The farmer chooses values of $x_j$, $\theta_j$, and $e_j$ given values of $\mu_j^p$ and $p_j$ for both $j \in \{K, L\}$.

The choice of $\theta_j$ is straightforward if the farmer selects a positive value for $x_j$: he will select $\theta_j = \hat{\theta}_j$ to minimize square loss in expectation. Thus, the farmer will need to decide the amount of $x_j$ to use in production and whether to place effort in learning about $j$'s production process ($e_j$). Let $P$ represent expected payoffs.[29] The farmer's problem can now

[29] We abuse notation slightly in expression (11) where P refers to the expected payoff, and not the arg max of the expression.

be summarized as:

$$
\begin{aligned}
\max P = \quad & \max_{(x_j,e_j)_{j\in\{K,L\}}} w_0 + x_1(1-p_1) + \textstyle\sum_{j\in\{K,L\}}[x_j(\mu_j^p - p_j) - x_j^2\sigma_{\theta_j}^2(e_j) - \bar{e}_j 1(e_j = 1)] \\
& \text{such that} \\
(\lambda): \quad & w_0 - \textstyle\sum_{j\in\{K,L\}} p_j x_j \geq 0.
\end{aligned} \tag{11}
$$

**Solution** Given the discrete nature of learning effort in our setup, we can find the solution to problem (11) by backward induction. We first determine the optimal value of $x_j$ given each choice of $e_j$; and then plug this value back into the objective function of problem (11) to determine which levels of effort result in the highest payoff.

First order conditions on $x_j$ yield the following demand function:

$$
x_j^* = \left(\frac{\mu_j^p - p_j(1+\lambda)}{2\sigma_j^2(e_j)}\right) \tag{12}
$$

Note that, $x_j^*$ is determined by values of $\mu_j^p$ and $e_j$. Expression (12) intuitively shows that demand for production technology $x_j$ is increasing in the net (perceived) returns of the production method and knowledge of the production process (recall, increased knowledge indicates a smaller value for $\sigma_j^2(e_j)$). Demand is decreasing with the penalty associated with violating the budget constraint, $\lambda$. However, the presence of borrowing constraints will result in corner solutions (of $x_j$) and non-adoption. Depending on the exact nature of the constraint, and the farmer's belief regarding the expected profitability of technology $j$, the farmer may decide to limit their attention to only learning about one or the other technology. Specifically, the farmer chooses to allocate learning effort on a combination of technologies, resulting in four possible combinations: $\mathbf{e} = [\{0,0\},\{1,0\},\{0,1\},\{1,1\}]$. Notice that $P(x_K^*(e_K), x_L^*(e_L), e_L, e_K)$ can be calculated for each of the four discrete choices a farmer can make. Thus, the farmers optimal learning effort vector, $\mathbf{e}^*$, can be characterized by:

$$
\begin{aligned}
\mathbf{e}^* \text{ such that for every } \mathbf{e}' &\in [\{0,0\},\{1,0\},\{0,1\},\{1,1\}] \\
P(x_K^*(e_K^*), x_L^*(e_L^*), e_L^*, e_K^*) &\geq P(x_K^*(e_K'), x_L^*(e_L'), e_L', e_K').
\end{aligned} \tag{13}
$$

From (11) it is obvious that the farmer will not excert any effort in learning the new technologies if the yields (net of costs) are smaller than 1 - the yield of the traditional technology. The choice of effort, in all other circumstances will depend on the net benefit from learning about technology $j$ relative to alternative technologies as described in equation (13); i.e., the benefit that comes from decreasing uncertainty about the optimal input level net of the cost of learning.

Holding the price of inputs and cost of learning effort fixed, this would imply finding a positive relationship between posterior yield beliefs and knowledge with any given technology. In regression terms, we would estimate the following equation:

$$knowledge_{ij} = \alpha + \beta_1 \mu_{ij}^p + \varepsilon_{ij} \tag{14}$$

**Choosing what to learn** A simple association between posterior beliefs and knowledge may be insufficient to explaining the nuanced ways in which farmers choose to learn about new technologies. To analyze the solution further, we make simplifying assumptions to the parameters in the model. Recall that we assumed that the capital-intensive technology generate higher average returns, i.e. $\mu_K > \mu_L$. We now, in addition, assume that they are also more expensive to purchase, or: $p_K > p_L$. Furthermore, we denote $\pi_j$, the average profit gain from an added unit of technology $j$, (i.e., $\pi_j = \mu_j - p_j$) and we assume that $\pi_K > \pi_L$. Finally, we assume that the cost of learning about technology $L$ is the same as that of learning about technology $K$ and the knowledge-benefit is similarly equivalent: $\{\bar{e}_K, \sigma_K^2(0), \sigma_K^2(1)\} = \{\bar{e}_L, \sigma_L^2(0), \sigma_L^2(1)\} = \{\bar{e}, \sigma^2(0), \sigma^2(1)\}$.

When the budget constraint binds, then $\lambda > 0$, and (11) can be solved by entering optimal values of $x_j^*$ into the budget constraint and equating the left and right-hand sides. In other words, we obtain a solution for $\lambda$ when $w_0 - \sum_{j \in \{K,L\}} p_j x_j^* = 0$ by replacing expressions of $x_j^*$ with the expression in equation (12).

$$\lambda = \begin{cases} \frac{p_L \pi_L \sigma^2(e_L) + p_K \pi_K \sigma^2(e_K) - 2w_0 \sigma^2(e_L)\sigma^2(e_K)}{p_L^2 \sigma^2(e_K) + p_K^2 \sigma^2(e_L)} & \text{if } w_0 = \mathbf{px}, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

As can be seen, the Lagrangian multiplier, or borrowing-constraint penalty, $\lambda$, is decreasing in wealth and exhibits a complex relationship between input price and knowledge. Specifically, there is a cross-technology knowledge-uncertainy trade-off that manifests itself in the multiplication of the knowledge-penalty of one technology with the price of the second technology. Depending on the underlying parameter space, this trade-off will lead to selective learning about one technology over the other if the wealth constraint is binding.

We can now compute the expected payoffs for the optimal solution using equation (11). In particular, when borrowing constraints do not bind, then $\lambda = 0$ and $x_j$ can be computed for all combinations of learning effort using equation (12). Plugging this information back into (11)) will, by comparing across the four alternatives, yield the optimal combination of effort and technology uptake; and resulting expected payoff.

When borrowing constraints do bind, then $\lambda$ is given by equation (15) and we can similarly compute expected payoff for all combinations of learning effort. Then, for any parameter combination, we compare expected payoffs across the four alternatives and identify $P^*$, as defined by equation (13), as the maximal value across the alternatives.

Figure 1 shows the relationship between the optimal expected payoff, which we denoted by $\pi$, and the farmer's initial wealth $w_0$ for each of the four possible learning combinations. We graph each learning combination separately. The largest expected payoff is determined by the particular value of initial wealth each farmer holds. Notice that expected payoffs are monotonically increasing in wealth but that there are thresholds at which farmers may choose to learn about neither technology $K$ or $L$ (lowest wealth category), either one of $K$ or $L$ (mid-tier wealth), or both $K$ and $L$ (unconstrained by wealth). Thus, we should only expect wealthy farmers to learn about the most capital-intensive components of new technologies. However, this is strongly contingent on the assumption that $\pi_k >> \pi_{L.}$ If beliefs about $\pi_k$ are not sufficiently high, then even a wealthy farmer will choose not to learn about technology $K$ because it is preferable to specialize in the labor-intensive mode of production.[30]

The regression implications of Figure 1 could be captured by:

$$knowledge_{iK} = \alpha_K + \beta_{1K}\pi_{iK} + \beta_{2K}w_i + \beta_{3K}w_i\pi_{iK} + \varepsilon_{iK}, \tag{16}$$

where $\pi_{iK}$ captures the profit gain observed by farmer $i$ and $w_i$ captures farmer $i$'s wealth. $\beta_{3j}$ is our coefficient of interest: If farmers have no reason to believe that an capital-intensive method of production will generate higher profits, then they will not choose to learn about this method of production. However, if farmers believe the new method will be profitable, we only expect learning to take place when farmers possess sufficient wealth. Thus, we hypothesize that $\beta_{3K}$ will be positive.

# Analysis of the learning process

In this final section, we present the results of the regression specifications presented in the previous section, further documenting the learning process.

## Correlates of yield expectations

In Table 6 we estimate regression specification (7) and regress endline yield expectations for soybean, hybrid maize and local maize (in kg/ha) on yield expectations at baseline (in kg/ha)

[30] This is demonstrated in Appendix Figure 4, which relaxes this assumption and shows that farmers will varies the difference between $\pi_K$ and $\pi_L$ while holding $w_0$ fixed at a sufficiently high level (allowing adoption of the capital-intensive technology). Notice that the farmer will never choose to learn about $K$ when $\pi_K - \pi_L$ is sufficiently small — and certainly will never learn about $K$ when $\pi_k < \pi_L$.

and the performance of the local demonstration plots (in particular, the mean differences between BPA and control subplots; also in kg/ha, adjusted for the moisture content) and a series of control variables. We split the sample according to absolute difference in soil quality between the local demonstration plot and the farmer's plot (measured by active carbon in mg/kg): Columns (1), (3) and (5) include the observations below the median of this distribution ("similar"), while Columns (2), (4) and (6) include the observations above the median of this distribution ("different") (the median absolute difference is around 125mg/kg for the soy demo sites, and 146 mg/kg for the maize demo sites). Note that we only included the CDI club farmers in the demonstration plot villages in these analyses.[31]

The results indicate that an increase of 1 kg/ha in the gap between the BPA and control soybean subplots is associated with an increase in the endline beliefs of 0.77 units (out of a possible score of 20) for farmers who have soils similar the demonstration plot soil, and is not associated with an increase (or decrease) for farmers whose soils are more dissimilar. The results for hybrid maize are qualitatively similar. An increase of the BPA performance on the local demonstration plot (compared to the control plots) by one kg/ha is associated with an increase in endline beliefs of 1.00 (significant at the 1 percent level). There is no statistically significant effects for farmers whose soil is dissimilar; nor is there any effect on the yield expectations of local maize. The latter can be interpreted as a placebo test: The CDI demonstrations did not include local maize, and hence nothing could have been learned about this crop.

## Correlates of knowledge

In Table 7 we estimate regression specification (14). We use demonstration plot yields, rather than self-reported beliefs, to capture the farmer's beliefs about profitability. We plot the relationship between demonstration plot yields and learning, as measured through the crop-specific knowledge score for the demonstration plot participants (again, including

[31] We focus on analysis of the CDI farmers in demonstration plot villages here for three reasons: (i) There is little variation in the demonstration plot performance observed by the field-day participants as they attended only one of two field-day sites (and we collected data only at one site which was within our study area), (ii) We did not collect soil data among all farmers in the field-day villages, and (iii) We noted that the quality of the yield expectation data was poorer in the non-demonstration plot villages. We attribute this lack of accuracy to the fact that farmers were asked to imagine the yield on a plot of one acre. As noted earlier, for some farmers, one acre is difficult to imagine. Farmers who attended the demonstration plot, having measured out the demonstration plot, and also having had one of their own plots measured, appeared to have an easier time with this type of question. Still, in interpreting the results it is important to recall that yield expectation question asked did not explicitly refer to median, average, or another moment - and hence the interpretation was up to the respondent.

a series of control variables). We find a statistically significant, and positive, relationship between the yields measured, and the average learning of the demonstration plot participants (soybean is almost statistically significant at the 10 percent level) - in Columns (1) and (3) for both soybean and hybrid maize. The magnitude indicates that an increase in 50 kg/ha increases the knowledge by approximately 5 percent (relative to the average score).

To interrogate the causal interpretation of the results, Columns (2) and (4) present the results using attributes of the demonstration plot rainfall rather than plot yields. Indeed reverse causality could explain the results presented in Columns (1) and (3): farmers with better knowledge have better results. Rainfall, on the other hand, does not suffer this critique and is correlated with maize development (Bradfort 1990, Çakir 2004, Lobell et al. 2011).

In Appendix Figure 3 we plot the rainfall distributions for demonstration plots. The rainfall distribution is quite variable. We defined three statistics of the distribution: start of the rainy season, the total amount of rainfall, and the number of flood days (defined as > 50 mm/day). The latter, in particular at the start of the season, can be quite damaging for germination (Wenkert et al. 1981, Martin 1991, Githiri et al. 2006). We note the anticipated correlation between rainfall patterns and knowledge of soy (but no such correlation for hybrid maize, which suggests that either rainfall and germination rates were not as closely related, or farmers' learning was more uniform across the various plots, perhaps because maize is a historically important crop).

In Table 8 we present the relationship between learning, also measured using the knowledge score, and yield expectations at endline for farmers were invited to participate in field-days. As in previous analyses, we split the knowledge score into knowledge related to soybean and knowledge related to maize. But this time, based on discussion of Figure 1, we include only the credit-intensive technologies in the soybean score, and the labor-intensive technologies in the maize score. We focus on one control variable: farmer wealth. We use "having obtained input credit in the previous season" as a proxy for the relevant wealth variable. If the farmer answers no to this question, the farmer is likely more credit constrained. We note a positive correlation between the endline beliefs and the knowledge score for soybean. The coefficient on our credit-measure suggests that farmers who are not credit constrained are more likely to learn something about the (credit-intensive) soybean technologies. The (hybrid) maize results act as a placebo test. The maize questions in the knowledge test, and included here, focused on labor-intensive technologies; and here we find no distinction in learning between credit constrained and non credit constrained farmers.

Finally, we consider a non-parametric version of equation (16) in Figure 2. On the left-hand side of the panel in this Figure are farmers who observed a small difference in yields between the soybean BPA and control plots (greater than the median) while the right-hand side are farmers who observed relatively large differences. The horizontal axis represents farmer wealth levels using a continuous asset index and the vertical axis represents farmers' knowledge of soybean best practice conditional on a set of individual characteristics.[32] According to our hypothesis, only wealthy farmers on the right-hand side panel should exhibit higher levels of learning, which we are able to confirm.

Note that Figure 2 considers soybean only. This relies on the assumption that soybean production is more capital-intensive and, therefore, more expensive. However, as a cash crop, it also potentially generates higher profits. We do not expect knowledge scores for hybrid maize to change with wealth since its production does not require as costly of inputs as soybean production under CDI's demonstrated best-practices (and effectively, this is what we find in Table 8).

It is notable that credit and wealth matter both for field-days and demonstration plots. Field-day participants likely face a significantly higher cost of learning (within that day); and for them, the trade-off between which technology to focus on likely much more salient. Focus group participants noted the rushed nature of the field-day and 'information overload' repeatedly. But even for the demonstration plot participants, wealth appears to matter, and learning might be concentrated among the wealthy farmers motivated by observed demonstration plot success.

# Conclusion

We studied farmers' learning about agricultural technologies based on differential exposure to commonly used extension methods that range in their intensity of interaction. We find that farmers who participated in farmer-led demonstration plots learn about the production processes of ISFM technologies critical for actual adoption including the type and amount of pesticide to be used on soy. Farmers who participated in farmer-led demonstration plots form beliefs about the usefulness of the technologies (as reflected in their yield expectations), with beliefs conditional on both the agronomic performance of the demonstration plot and how similar their own soil conditions are to the demonstration plot soil conditions. These elements, in their turn, correlate with the formation of knowledge about the production processes. Not surprisingly, farmers invited to attend field-days learn considerably less about the production processes than those involved in managing demonstration plots. However, we find that what farmers learn is conditional on the degree to which they are credit-constrained.

[32] We apply a control-function approach to variable construction by using the predicted error term on the vertical access after regressing farmer knowledge against gender, age, years of education, household size (total and adult), land size and wealth.

This result is corroborated by qualitative interviews suggesting that field-day participants tend to focus their attention on learning labor-intensive technologies, such as plant spacing and mulching.

These results suggest the presence of a two-stage learning process. In the first stage, farmers form yield expectations, i.e., conditional on applying the ISFM technologies, what would my maize yield be? In the second stage, farmers choose to invest in learning the details of the production processes subject to these perceived yield beliefs and the cost of learning.

Differences in how this two-stage learning process develops between demonstration plots and field-days are reflected in farmer adoption plans. We find that only farmers participating in demonstration plots are more likely to adopt hybrid maize, inoculate soy and plant fertilizer trees. Other researchers have also reported positive effects of demonstration plots on adoption, especially among farmers directly involved in plot management, such as, Duflo et al. (2006), Kondylis et al. (2017) and Lunduka et al. (2018). Farmers invited to participate in field-days, on the other hand, do not, on average, adopt more ISFM technologies.

This result suggests that farmer field-days in agro-climatic zones different from farmers' own might not result in widespread adoption of a new technology, especially if farmers need to be convinced that the technology will increase average yields *prior* to investing in learning the production processes. If, in addition, a field-day is too short to learn all of the production processes, farmers might not be able to easily progress to this second stage of learning, even if convinced about the technology's yield-increasing attributes.[33]

Our results have implications for Malawi and other Sub-Saharan African countries working to reform extension systems, moving from traditional training and visit systems to systems which aim to be more demand-driven, accountable and cost-effective (for a discussion on the status of Malawi's extension system, see MEAS 2012 and MAIWD 2016). This change is the result of a combination of declining budgets for extension, partially due to a donor retreat from extension, and the perceived lack of effectiveness of traditional extension systems (see, for instance, Davis 2008, Evenson 1997, Anderson et al. 2006). The Malawian government's extension system is under significant strain, under-resourced and under-incentivised. Extension workers, generally equipped with a bicycle only, are expected to cover long distances and to conduct a range of government and non-government activities with minimal

[33] This distinction between knowing about the existence of the technology and learning its attributes has also been documented by others. Kabunga et al. (2012) noted that while many farmers in Kenya have heard about tissue culture in bananas, few know the details required to implement the technology. Lambrecht et al. (2014) find that while awareness about fertilizers has spread widely among farmers in Congo, direct contact with extension agents is what contributes to adoption.

support (see Knorr et al. 2017).[34] Many institutions for training agricultural extension agents have closed and those remaining now require that students pay their own fees for a MS degree in extension. In turn, extension workers receive a comparatively small, monthly, fixed salary. This situation has reportedly led to pervasive problems with moral hazard and adverse selection. The type of student that can afford the requisite training might not be interested in returning to work in rural areas and ends up being employed by NGOs or other private institutions, draining the government extension system. The fixed salary might further reduce the uncontractabl) effort of extension officers (MEAS 2012, CISANET 2013, MAIWD 2016). BenYishay and Mobarak (2013), ushow that incentivizing extension agents in Malawi by paying them for farmer knowledge improvements improves farmers' uptake of these technologies. Alternatively, one might increase the strength of the stick, as in Dal Bo et al. (2018) who show that tracking the extension workers via GPS (in Paraguay), especially when the ones to be tracked are chosen by the supervisors, might increase the effectiveness of extension services.

Despite these challenges, government extension workers are still a main source of information for farmers. Ragasa and Niu (2017), in a comprehensive overview of access and demand for agricultural information, note that almost 70 percent of households who received advice from external sources received it from government extension agents. In our study area in central Malawi, each government extension agent is in charge of 2,000 to 3,000 farming households.[35]

Our results have implications for such an extension system and suggest that field-days should be carefully designed and that extension programs should have the right expectations about what such efforts might accomplish. We find that farmers learn more from demonstration plots compared to farmer field-days, not surprising given the time farmers spend on a demo plot versus a field day. However, field-days are often a less expensive option and serve as a common cost-saving strategy of extension plans. A back-of-the-envelope calculation demonstrates that the per-farmer cost difference is substantial between a field day and a demonstration plot; in our study, hosting one farmer-field day for around 200 farmers costs about 650 USD total, or about 3.25 USD per farmer (we used the number of field-day

[34] The latter has been challenged: Niu and Ragasa (2017) assess information efficiency along the knowledge transmission chain from researchers to extension agents, lead farmers, and other farmers. They asked all parties questions about pit-planting, and use these to construct a measure of knowledge at each node of the knowledge transmission chain. They find that the majority of information loss happens at the extension agent to lead farmer link.

[35] However, as the distance between the villages covered and the homestead of the extension agent can vary substantially, a heterogenous coverage is to be expected, and only 60 percent of the farmers in our study report interaction with their government extension agent in the last year.

participants in Mtumthama for this estimate), while organizing one demonstration plot for about 20 farmers costs 281 USD, about 14 USD per farmer.

We re-iterate that this study is not an evaluation of the relative effectiveness of field-days versus demonstration plots and our results do not suggest that field-days should be discarded as a strategy. We make suggestions below building on what we have learned about the learning process to suggest improvements in the field-days.

First, farmer field-days may provide too much information in too short a time period, giving farmers insufficient chance to absorb the details. This implies that, at field-days, farmers should be given tools which will allow them to learn the information presented more effectively. Examples might include pamphlets with pictures of the inputs used and measuring spoons to measure the correct amounts of inputs.[36]

Second, the fact that farmers' learning appears to be constrained by markets suggests that agricultural extension might need a re-coupling with market activities, and in particular, credit interventions in order to be effective. In Malawi, extension agents used to perform an additional role as regional credit officers. While conflict of interest should be avoided, providing farmers access to credit while introducing a new intervention is likely to affect uptake and learning given evidence that credit access itself influences how open the farmer is to receiving information on capital-intensive technologies. Ambler et al. (2017), using a randomized control trial set in Malawi, also note complementarities between cash transfers and a more intensive (as opposed to extensive, standard) extension system. However, Ragassa and Mazunda (2018) find no effect of the interaction effect of access to Malawi's input subsidy program and having received advice from extension agents (but attribute some of the lack of effects to heterogeneities within the extension program).

Third, heterogenous growing conditions may play a role in influencing what farmers take away from field-days. In this regard, an in-village demonstration plot might be a better choice, with the caveat that a bad yield could result in a "non-adoption" trap. Demonstration plots in various conditions are to be recommended; with participants being matched to attend field-days at demonstration plots that match their own growing conditions.

Fourth, while cell phone penetration is not as high in Malawi as it is in other Sub-Saharan African countries, using cell phone as a medium might provide an appealing alternative (for an overview on ICT based extension systems, see Aker 2011 and Davis 2008). But even in this case, SMS messages would need to use simple language, be growing-condition specific and might need to be accompanied by market information. SMS messages can also be utilised to repeat information, as in Dzanku et al. (2018), who reminded farmers of optimal storage practices at the time of harvest in Ghana.

[36] Duflo et al. (2013) find a large demand for fertilizer measuring spoons in Kenya.

Finally, it may be that field-days could be used in sequence with demonstration plots or other more intensive methods of teaching farmers. The field-days could serve to introduce a new technology and to focus on its broad features, demands, and processes and this initial introduction could be followed by methods employing more detailed exposure, perhaps based on farmer demand.

We conclude with a note on further research. While the limited time frame of this study (i.e., one year) does not allow for a detailed study of learning spill-overs (and, relatedly, strategic learning interactions between farmers), we recognize their importance and appeal to extension models: In the training and visit extension model, for example, extension agents are updated with the latest technologies and generally visit selected lead farmers who are expected to teach farmers in their community. The extent to which adoption spreads through the communities through social learning is expected to depend on the degree of heterogeneity between farmers, as well as the structure of the social network, the identities of the first adopters and whether and how lead farmers are encouraged (Griliches 1957, Foster and Rosenzweig 1995, Munshi 2004, Bandiera and Rasul 2006, Conley and Udry 2010, Chuang and Schechter 2015, Maertens 2017, Michelson 2017). Beaman et al. (2013) use a network-theory approach to better identify these lead farmers in order to maximize learning and adoption in their communities. Shikuku et al. (2017) and BenYishay and Mobarak (2013) both use a randomized control trial to vary incentives for, respectively, lead farmers and extension agents, and find that both respond to incentives. We see this type of research, which combines network theory with realistic models of learning and behavior with heterogenous agents (in terms of cognitive ability as in Barham et al. 2017, more general, skills as in Laajaj and Macours 2017A, or "locus of control", as in Malacarna 2018) as a fruitful way forward in extension research.

# References


[1] Abadie, A. and G.W. Imbens. 2009. "Matching on the Estimated Propensity Score." NBER Working Paper No. 15301.

[2] AGRA and IIRC. 2014. *Investing in soil: Cases and lessons from AGRA's Soil Health Programme*. Alliance for a Green Revolution in Africa and International Institute of Rural Reconstruction, Nairobi.

[3] Aker, J.C., 2011. "Dial "A" for agriculture: A Review of Information and Communication Technologies for Agricultural Extension in Developing Countries." *Agricultural Economics*, 42(6), pp.631-647.

[4] Ambler, K., de Brauw, A. and S. Godlonton. 2017. "Relaxing Constraints for Family Farmers: Providing Capital and Information in Malawi." Working Paper Presented at NEUDC 2017 at Tufst University.

[5] Anderson, J.R. and Feder, G., 2007. "Agricultural Extension." in: *Handbook of Agricultural Economics*, 3, pp.2343-2378.

[6] Anderson, J.R.,Feder, G. and Ganguly, S. 2006. "The rise and fall of Training and Visit extension: An Asia mini-drama with an African Epilogue," in A.W. Van den Ban and R.K Samantha (eds.) Changing Role of Agricultural Extension in Asian nations. new Delhi: B.R Publishing Corporation, pp. 149-74.

[7] Bandiera, O. and Rasul, I., 2006. "Social Networks and Technology Adoption in Northern Mozambique." *The Economic Journal*, 116(514), pp.869-902.

[8] Barham,B.L., Chavas,J.P., Fitz, D. and L. Schechter. 2017. "Receptiveness to Advice, Cognitive Ability, and Technology Adoption: An Economic Experiment with Farmers." Working Paper.

[9] Beaman, L., Karlan, D., Thuysbaert, B. and Udry, C., 2013. "Profitability of Fertilizer: Experimental Evidence from Female Rice Farmers in Mali." *The American Economic Review*, 103(3), pp.381-386.

[10] Beaman, L., BenYishay, A., Magruder, J., and A.M. Mobarak. 2013. "Can Network Theory-based Targeting Increase Technology Adoption?" Working Paper.

[11] BenYishay, A. and Mobarak, A.M., 2013. "Communicating with Farmers through Social Networks." Working Paper.

[12] Bevis, L.E., and C.B. Barrett, C.B. 2016. "Close to the Edge: Do Behavioral Explanations Account for the Inverse Productivity Relationship? Working Paper.

[13] Bezu, S., Kassie, G.T., Shiferaw, B. and Ricker-Gilbert, J., 2014. "Impact of Improved Maize Adoption on Welfare of Farm Households in Malawi: A Panel Data Analysis." *World Development*, 59, pp.120-131.

[14] Birkhaeuser, D., Evenson, R.E. and Feder, G., 1991. "The Economic Impact of Agricultural Extension: A Review." *Economic Development and Cultural Change*, 39(3), pp.607-650.

[15] Birner, R., Davis, K., Pender, J., Nkonya, E., Anandajayasekeram, P., Ekboir, J., Mbabu, A., Spielman, D., Horna, D., Benin, S. and Cohen, M. "From Best Practice to Best Fit: A Framework for Designing and Analyzing Pluralistic Agricultural Advisory Services Worldwide." The Journal of Agricultural Education and Extension 15(4): 341-355.

[16] Bradfort, Kent J. 1990. "A Water Relations Analysis of Seed Germination Rates," *Plant Physiology*, 94(2), pp.840-849

[17] Brown, P.C., Roediger, H.L. and McDaniel, M.A., 2014. *Make it Stick*. Harvard University Press.

[18] Bulte, E., Beekman, G., Di Falco, S., Hella, J. and Lei, P., 2014. "Behavioral Responses and the Impact of new Agricultural Technologies: Evidence from a Double-blind Field Experiment in Tanzania." *American Journal of Agricultural Economics*, 96(3), pp.813-830.

[19] Cakir, R. 2004. "Effect of Water Stress at Different Development Stages on Vegetative and Reproductive Growth of Corn." *Field Crops Research*, 89(1) pp: 1-16

[20] Chinangwa, L.L.R., 2006. Adoption of Soil Fertility Improvement Technologies among Smallholder Farmers in Southern Malawi. Master's Thesis, Department of International Environment and Development Studies, Norwegian University of Life Sciences.

[21] Chuang, Y. and Schechter, L., 2015. "Social Networks in Developing Countries." *Annual Review of Resource Economics*, 7(1), pp.451-472.

[22] Chafutsa, W., Mtukuso, A. and M. Banda. 2013. A guide to soybean production in Malawi. Department of Agricultural Research Services DARS, Malawi.

[23] Cai, Tian & Steinfield, Charles & Olson, Jennifer. (2017). Keeping Top-of-Mind: The Impact of Audio Phone Reminders on Kenya Farmers' Knowledge and Uptake of Drought Tolerant (DT) Maize. Presented at the AAEA Annual Meetings in Chicago.

[24] CISANET (Civil Society Agricultural Network). 2013. The State of Agricultural Extension Services in Malawi. Available from: http://www.cisanetmw.org/index.php/publications

[25] Conley, T.G. and Udry, C.R., 2010. "Learning about a New Technology: Pineapple in Ghana." *The American Economic Review*, 100(1), pp.35-69.

[26] Crane-Droesch, A. 2018. "Technology Diffusion, Outcome Variability and Social Learning: Evidence from a Field Experiment in Kenya." *American Journal of Agricultural Economics.*

[27] Davis, K., 2008. "Extension in sub-Saharan Africa: Overview and Assessment of Past and Current Models and Future Prospects." *Journal of International Agricultural and Extension Education*, 15(3), pp.15-28.

[28] Delavande, A., Giné, X. and McKenzie, D., 2011A. "Eliciting Probabilistic Expectations with Visual Aids in Developing Countries: How Sensitive are Answers to Variations in Elicitation Design?" *Journal of Applied Econometrics*, 26(3), pp.479-497.

[29] Delavande, A., Giné, X., and D. McKenzie. 2011B. "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence." *Journal of Development Economics*, 94 (2): 151-163.

[30] Desiere, S., and D. Jolliffe. 2018. "Land Productivity and Plot Size: Is Measurement Error Driving the Inverse Relationship." Journal of Development Economics 130, pp. 84–98.

[31] Dillon, B., 2015. "Measuring Subjective Probability Distributions." Working Paper.

[32] Dal Bo, E., F. Finan, N. Y., Li and L. Schechter. 2018. "Government Decentralization Under Changing State Capacity: Experimental Evidence from Paraguay." Working Paper.

[33] Droppelmann, K.J., S.S. Snapp, and S.R. Waddington. 2017. Sustainable intensification options for smallholder maze-based farming systems in sub-Saharan Africa. Food Security 9:133-150,

[34] Duflo, E., Kremer, M. and Robinson, J., 2008. "How High are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya." *The American Economic Review*, 98(2), pp.482-488.

[35] Duflo, E., Kremer, M., Robinson, J. and F. Schilbach, 2013. "Technology Diffusion and Appropriate Use: Evidence from Western Kenya." Working Paper.

[36] Dzanku, F. M, Osei, R. D, Osei-Akoto, I., Hodley, L. S., Adu, P.N. and K. Adu-Ababio. 2018. "Do Mobile Phone Voice Messages Reminders Following Face-to-Face Training Improve Adoption and Smallholder Farmer Outcomes." Presented at CSAE 2018.

[37] Emerick, K., de Janvry, A., Sadoulet, E. and Dar, M.H., 2016. "Technological Innovations, Downside Risk, and the Modernization of Agriculture." *The American Economic Review*, 106(6), pp.1537-1561.

[38] Evenson, R., 1997. "The Economic Contributions of Agricultural Extension to Agricultural and Rural Development." In: *Improving Agricultural Extension. A Reference Manual.* Swanson, B.E. (ed.) Bentz, R.P. (ed.) Sofranko, A.J. (ed.) / FAO, Rome (Italy). Research, Extension and Training Div. , p. 27-36.

[39] Fairhurst, T., 2012. *Handbook for Integrated Soil Fertility Management.* Africa Soil Health Consortium. Nairobi.

[40] Feder, G.,Willet, A., and Zijp,W., 1999. "Agricultural Extension: Generic Challenges and some Ingredients for Solutions." World Bank Policy Research Working Paper 2129, World Bank, Washington, DC.

[41] Fehr, E. and Rangel, A., 2011. "Neuroeconomic Foundations of Economic Choice—Recent Advances." *The Journal of Economic Perspectives*, 25(4), pp.3-30.

[42] Foster, A.D. and Rosenzweig, M.R., 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of political Economy*, 103(6), pp.1176-1209.

[43] Franke, A.C., Van Den Brand, G.J. and Giller, K.E., 2014. "Which Farmers Benefit Most from Sustainable Intensification? An ex-ante Impact Assessment of Expanding Grain Legume Production in Malawi." *European Journal of Agronomy*, 58, pp.28-38.

[44] Gabaix, X., Laibson, D., Moloche, G. and Weinberg, S., 2006. "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model." *The American Economic Review*, 96(4), pp.1043-1068.

[45] Gabaix, X. 2017. "Behavioral Inattention." NBER Working Paper 24096.

[46] Gatere, L. 2013. "Field Kit Soil Tests to Assess Acidity, N, P, S and K Fertility in Kenyan Soils," presented at American Society of Agronomy, Crop Science Society of America, and the Soil Science Society of America Annual Meetings in Tampa, Florida.

[47] Githiri, S. M., Watanabe, S., Harada K, and R. Takahashi. 2006. "QTL Analysis of Flooding Tolerance in Soybean at an Early Vegetative Growth Stage." *Plant Breeding* 125(6), pp. 613-618.

[48] Ghosh, S. 2016. "Costly Social Learning and Rational Inattention." Working Paper NYU.

[49] Griliches, Z., 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica*, 25(4), pp.501-522.

[50] Hanna, R., Mullainathan, S. and Schwartzstein, J., 2014. "Learning Through Noticing: Theory and Evidence from a Field Experiment." *The Quarterly Journal of Economics*, 129(3), pp. 1311–1353.

[51] Harstad, R. M. and R. Selten. 2013. "Bounded-Rationality Models: Tasks to become Intellectually Competitive." *Journal of Economic Literature,* 51(2), pp. 496-511.

[52] Hopper, W. D., 1965. "Allocation Efficiency in a Traditional Indian Agriculture." *American Journal of Agricultural Economics*, 47(3), pp. 611-624.

[53] Horowitz, J. L. 2001. "The Bootstrap," in the *Handbook of Econometrics*, vol. 5, pp 3159-3228.

[54] Independent Evaluation Group (IEG). 2011. *Impact Evaluations in Agriculture: An Assessment of the Evidence.* Washington, DC: World Bank.

[55] Kabunga, N.S., Dubois, T. and Qaim, M., 2012. "Heterogeneous Information Exposure and Technology Adoption: The Case of Tissue Culture Bananas in Kenya." *Agricultural Economics*, 43(5), pp.473-486.

[56] Kahneman, D., 1973. *Attention and Effort.* Prentice-Hall Series in Experimental Psychology. Englewood Cliffs, NJ: Prentice-Hall.

[57] Kahneman, D. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review*, 93 (5), pp. 1449-1475.

[58] Kassie, M., Marenya, P., Tessema, Y., Jaleta, M., Zeng, D., Erenstein, O. and D. Rahut. Forthcoming. "Measuring Farm and Market Level Economic Impacts of Improved Maize Production Technologies in Ethiopia: Evidence from Panel Data." *Journal of Agricultural Economics.*

[59] Kerr, R.B., Snapp, S., Chirwa, M., Shumba, L. and Msachi, R., 2007. "Participatory Research on Legume Diversification with Malawian Smallholder Farmers for Improved Human Nutrition and Soil Fertility." Experimental Agriculture, 43(04), pp.437-453.

[60] Kim, J.W, Ritter, F.E., and R.J. Koubek. 2013. "An Integrated Theory for Improved Skill Acquisition and Retention in Three Stages of Learning." *Theoretical Issues in Ergonomics Sciences*, 14(1), pp. 22-37.

[61] Knorr, J., Gerster-Betaya, M. and V. Hoffman. 2007. *The History of Agricultural Extension in Malawi.* Margraf Publishers GMbH, Scientific books.

[62] Kondylis, F., Mueller, V. and J. Zhu, 2017 "Seeing is believing? Evidence from an extension network experiment," *Journal of Development Economics,* 125, pp. 1–20.

[63] Kondylis, F., Mueller, V. and J. Zhu. 2015. "Measuring Agricultural Knowledge and Adoption." *Agricultural Economics*, 46:449-462.

[64] Krishnan, P. and Patnam, M., 2012. Neighbours and Extension Agents in Ethiopia: Who Matters more for Technology Diffusion? Working Paper

[65] Krueger, R.A. and Casey, M.A. 2010. "Focus Group Interviews." in: Handbook of Practical Program Evaluation. Evaluation, 3rd Edition, (eds.) Joseph S. Wholey, Harry P. Hatry, Kathryn E. Newcomer. San Fransisco, Josset-Bass, a Wiley Imprint.

[66] Laajaj, R. and Macours, K. 2017A. "Measuring Skills in Developing Countries." Working Paper.

[67] Laajaj, R. and Macours, K. 2017B. "Learning-by-Doing and Learning-from-Others: Evidence from Agronomical Trials in Kenya" Working Paper.

[68] Lichand, G. and A. Mani, 2017. "Cognitive Droughts." Working Paper.

[69] Lipton, M and R Longhurst, 1989. *New Seeds and Poor People*. Routledge Library Editions: Development.

[70] Lybbert, T.J., Barrett, C.B., McPeak, J. G., and Winnie K. Luseno. 2007. “Bayesian Herders: Updating of Rainfall Beliefs In Response to External Forecasts,” *World Development* 35, pp. 480-497.

[71] Lunduka, R.W., Snapp, S., and T.S. Jayne. 2018."Demand-Led and Supply-Led Extension Approaches to Support Sustainable Intensification in Malawi." Working Paper.

[72] Lobell, D.B, Bänziger, M., Magorokosho, C and V. Bindiganavile. 2011. "Nonlinear Heat Effects on African maize as Evidenced by Historical Yield Trials." *Nature Climate Change*, 1, pp.42–45.

[73] Maertens, A., 2017. "Who Cares what Others Think (or do)? Social Learning and Social Pressures in Cotton Farming in India." *American Journal of Agricultural Economics.*

[74] Malacarna, J.G., 2018. "Locus of Control and Investment Behavior Among Smallholding Maize Farmers: An Empirical Study from Mozambique and Tanzania." Presented at PACDEV, UCDavis.

[75] Manda, J., Alene, A.D., Gardebroek, C., Kassie, M. and Tembo, G., 2016. "Adoption and Impacts of Sustainable Agricultural Practices on Maize Yields and Incomes: Evidence from Rural Zambia." J*ournal of Agricultural Economics*, 67(1), pp.130-153.

[76] Marenya, P.P. and C.B. Barrett. 2007. "Household-Level Determinants of Adoption of Improved Natural Resources Management Practices among Smallholder Farmers in Western Kenya," *Food Policy*, 32(4), pp. 536.

[77] Marenya, P.P. and Barrett, C.B., 2009. "State-Conditional Fertilizer Yield Response on Western Kenyan Farms." *American Journal of Agricultural Economics*, 91(4), pp.991-1006.

[78] MEAS (Modernizing Extension and Advisory Services). 2012. "Strenthening Pluralistic Agricultural Extension in Malawi." Report on the MEAS Rapid Scoping Mission carried out January 7-27, 2012.

[79] Martin, B. A., S. F. Cerwick, and L. D. Reding. 1991. "Physiology Basis for Inhibition of Maize Seed Germination by Flooding." *Crop science* 31(4), pp. 1052-1057.

[80] Mhango, W.G., Snapp, S.S. and Phiri, G.Y., 2013." Opportunities and Constraints to Legume Diversification for Sustainable Maize Production on Smallholder Farms in Malawi." *Renewable Agriculture and Food Systems*, 28(03), pp.234-244.

[81] Michelson, H. C. (2017). " Influence of Neighbor Experience and Exit on Small Farmer Market Participation." *American Journal of Agricultural Economics*, 99(4), pp.952-970.

[82] MAIWD (Ministry of Agriculture, Irrigation and Water Development), Govt. of Malawi. 2016. National Agricultural Policy.

[83] MAIFS (Ministry of Agriculture, Irrigation and Food Security), Govt. of Malawi. 2012. Guide to Agriculture Production and Natural Resource Management in Malawi.

[84] Morgan, D.L., 1996. *Focus Groups as Qualitative Research*. Series on Qualitative Research Methods (Vol. 16). Sage Publications.

[85] Mponela, P., Tamene, L., Ndengu, G., Magreta, R., Kihara, J. and Mango, N., 2016. "Determinants of Integrated Soil Fertility Management Technologies Adoption by Smallholder Farmers in the Chinyanja Triangle of Southern Africa." *Land Use Policy*, 59, pp.38-48.

[86] Mugwe, J., Mugendi, D., Mucheru-Muna, M., Merckx, R., Chianu, J. and B. Vanlauwe. 2009. "Determinants of the Decision to Adopt Integrated Soil Fertility Management Practices by Smallholder Farmers in the central Highlands of Kenya." *Experimental Agriculture*, 45(1), pp. 61-75.

[87] Munthali, M.W., 2007. "Integrated Soil Fertility Management Technologies: A Counteract to Existing Milestone in Obtaining Achievable Economical Crop Yields in Cultivated Lands of Poor Smallholder Farmers in Malawi." in: Bationo A., Waswa B., Kihara J., Kimetu J. (eds.) *Advances in Integrated Soil Fertility Management in Sub-Saharan Africa: Challenges and Opportunities*, pp.531-536.

[88] Munshi, K., 2004. "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution." *Journal of Development Economics*, 73(1), pp.185-213.

[89] Niu, C. and C. Ragasa. 2017. "Limited attention and information loss in the lab-to-farm knowledge chain: The case of Malawian agricultural extension programs." IFPRI discussion paper, pp. 56.

[90] Njoloma, J.P., Sileshi, W.G., Sosola, B.G., Nalivata, P.C. and Nyoka, B.I., 2016. "Soil Fertility Status Under Smallholder Farmers Fields in Malawi." *African Journal of Agricultural Research*, 11(19), pp.1679-1687.

[91] Nkonya E., Johnson, T., Kwon, H.Y. and E. Kato. 2016. "Economics of Land Degradation in Sub-Saharan Africa," In: E. Nkonya, A. Mirzabaev and J. von Braun (eds). *Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development.* Springer, pp. 215-260.

[92] Nkonya, E., Azzari, C., Kato, E., Koo, J., Nziguheba, G. and B. Vanlauwe. 2017. *Mapping Adoption of ISFM Practices Study: The Case of Keny,a Rwanda and Zambia.* IFPRI AND IITA Report.

[93] Nourani, V. 2016. "Social Network Effects of Technology Adoption: Investing with Family, Learning from Friends and Reacting to Acquaintances". Working paper.

[94] Ortega, D.L., Waldman, K.B., Richardson, R.B., Clay, D.C. and Snapp, S., 2016. "Sustainable Intensification and Farmer Preferences for Crop System Attributes: Evidence from Malawi's Central and Southern Regions." *World Development*, 87, pp.139-151.

[95] Owens, T., Hoddinott, J. and Kinsey, B., 2003. "The Impact of Agricultural Extension on Farm Production in Resettlement Areas of Zimbabwe." *Economic Development and Cultural Change,* 51(2), pp.337-357.

[96] Pan, D., Ying, R., Zhang, N., and F. Kong. 2017. "Does Agricultural Training Improve Farmers' Fertilizer Use Knowledge? Evidence from a Randomized Controlled Experiment in Chinese Rice Production." Presented at the AAEA Annual Meetings in Chicago.

[97] Pan, Y., S.C. Smith and M. Sulaiman. Forthcoming. "Agricultural Extension and Technology Adoption for Food Security: Evidence from Uganda." *American Journal of Agricultural Economics.*

[98] Piciotti, R. and J.R. Anderson, 1997. "Reconsidering Agricultural Extension." *World Bank Research Observer*, 12(2), pp. 249-59.

[99] Place, F., Barrett, C,B, Freeman, H.A, Ramisch, J.J. and B. Vanlauwe, 2003. "Prospects for Integrated Soil Fertility Management using Organic and Inorganic Inputs: Evidence from Smallholder African Agricultural Systems." Food Policy, 28(4), pp. 365-378.

[100] Rabin, M. 2013. "Incorporating Limited Rationality into Economics." Journal of Economic Literature, 51(2), pp. 528-43.

[101] Ragasa, C. and C. Niu. 2017. *The state of agricultural extension and advisory services provision in Malawi: Insights from household and community surveys.* Report International Food Policy Research Institute (IFPRI).

[102] Ragasa, C. and J. Mazunda. 2018. "The Impact of Agricultural Extension Services in the Context of a Heavily Subsidized Input System: The Case of Malawi," *World Development*, 105, pp. 25-47.

[103] IHS, Republic of Malawi. 2017. Integrated Household Survey 2016-17: Household Socio-Economic Characteristics Report.

[104] Rogers E. M., 1995. *Diffusion of Innovations.* Free Press; New York.

[105] Rozenbaum, P.R., and P. Rubin, 1983. "The Central Role of the Propensity score in Observational Studies for Causal Effects." 70 (1), pp. 41-55.

[106] Sauer, J. and Tchale, H., 2009. "The Economics of Soil Fertility Management in Malawi." *Review of Agricultural Economics,* 31(3), pp.535-560.

[107] Sanchez, P.A. 2002. "Soil Fertility and Hunger in Africa." *Science*, 295 (5562), pp. 2019-2020.

[108] Shikuku, K.M., Pieters, J., Bulte, E., and P. Läderach. 2017. "Pro-Social Behavior in Agricultural Knowledge Diffusion." Presented at SPIA Conference, ICRAF, Nairobi.

[109] Schilbach, F., Schofield, H. and S.Mullainathan. 2016. "The Psychological Lives of the Poor." *American Economic Review*, 106(5): 435-40.

[110] Schultz, T.W. 1964. *Transforming Traditional Agriculture (Study in Comparative Economics).* New Haven: Yale University Press, pp. 226.

[111] Snapp, S.S. 1998. "Soil nutrient status of smallholder farms in Malawi." *Communication in Soil and Plant Analysis* 29 (17&18),pp. 2571-2588.

[112] Suri, T., 2011. "Selection and Comparative Advantage in Technology Adoption." *Econometrica*, 79(1), pp.159-209.

[113] Tjernström, E., 2015. "Signals, Similarity and Seeds: Social Learning in the Presence of Imperfect Information and Heterogeneity." Working Paper.

[114] Sheahan, M., and C. B. Barrett,2017. "Ten Striking Facts about Agricultural Input Use in Sub-Saharan Africa." *Food Policy*, 67, pp. 12-25.

[115] Spielman, D. 2017. "The Role of Evidence in Designing Innovative Approaches to Agricultural Extension & Rural Advisory Services." Presented at the 2017 AAEA Meetings in Chicago.

[116] Tully, K., Sullivan, C., Weil, R. and Sanchez, P., 2015. "The State of Soil Degradation in Sub-Saharan Africa: Baselines, Trajectories, and Solutions." *Sustainability*, 7(6), pp.6523-6552.

[117] Vanlauwe, B. and K.E. Giller, 2006. "Popular Myths around Soil Fertility Management in Sub-Saharan Africa." *Agriculture, Ecosystems and the Environment*, 116, pp 34–46.

[118] VanKehn, K. 1996. "Cognitive Skill Acquisition," *Annual Review of Psychology*, 47(1), pp. 513-415.

[119] Waddington, H. and H. White. 2014. *Farmer Field Schools: From Africultural Extension to Adult Education.* 3IE Systematic Review Summary.

[120] Weil, R.R, Islam, K.R, Stine, M.A., Gruver, J.B. and S.E. Samson-Liebig. 2003. "Estimating Active Carbon for Soil Quality Assessment: A Simplifed Method for Laboratory and Field Use." *American Journal of Alternative Agriculture*, 18(1), pp. 3-17.

[121] Weil, R. and L. Gatere. 2015. SoilDoc Kit System, BETA Version of SoilDoc Protocols. Manual. Agricultural and Food Security System. Earth Institute. Columbia University.

[122] Wenkert, W., Fausey, N.R. and D. Watters, H. 1981. "Flooding Responses in Zea mays L." *Plant and Soil*, 62(3), pp. 351-366.

[123] Wossen, T., Berger, T. and Di Falco, S., 2015. "Social Capital, Risk Preference and Adoption of Improved Farm Land Management Practices in Ethiopia." *Agricultural Economics*, 46(1), pp.81-97.

**Table 1: Descriptive statistics of households at baseline in 2014**

| Variable description | N | Mean | St. Dev. |
|---|---|---|---|
| *Panel A: Socio-economic characteristics* | | | |
| Gender of household head (0=male; 1=female) | 1,000 | 0.18 | 0.38 |
| Age of household head (years) | 1,000 | 42.45 | 15.01 |
| Education of household head (years of education) | 1,000 | 4.59 | 3.43 |
| Number of household members | 1,000 | 5.22 | 2.14 |
| Agriculture main activity of household head (1=yes; 0=no) | 1,000 | 0.79 | 0.41 |
| Land (in acres, owned)[1] | 944 | 3.47 | 2.55 |
| Are government extension agents one of three main sources of information (no=0; yes=1)[2] | 1,000 | 0.30 | 0.46 |
| Are other farmers one of three main sources of information (no=0; yes=1)[2] | 1,000 | 0.75 | 0.42 |
| Took credit in 2013-14 season(no=0; yes=1)[3] | 1,000 | 0.17 | 0.37 |
| *Panel B - Perceived soil quality* [4] | | | |
| Perceived stagnant or declining soil fertility (no=0; yes=1) | 960 | 0.82 | 0.34 |
| Experienced soil erosion (no=0; yes=1) | 960 | 0.47 | 0.44 |
| Experienced nutrient depletion (no=0; yes=1) | 960 | 0.57 | 0.45 |
| Experienced water logging (no=0; yes=1) | 960 | 0.23 | 0.37 |
| Experienced acidity or salinity (no=0; yes=1) | 960 | 0.05 | 0.20 |
| *Panel C - Results of soil sample analysis* [5] | | | |
| pH (recall 7 is neutral, smaller is acid, larger is alcalic) | 252 | 6.12 | 0.52 |
| Active carbon ( in mg/kg) | 250 | 423 | 150 |
| Limited N (no=0; yes=1) | 252 | 1.00 | 0.00 |
| Limited S (no=0; yes=1) | 252 | 0.77 | 0.43 |
| Limited K (no=0; yes=1) | 252 | 0.33 | 0.47 |
| Limited P (no=0; yes=1) | 252 | 0.52 | 0.55 |

Notes: 1: We dropped farmers with more than 13 acre (95% percentile) for these statistics. 2: We asked the respondent about the three main sources of information about agriculture, if the government extension agent was mentioned, we coded the first answer = yes (no otherwise); and if another farmer in the village was mentioned, we coded the second answer = yes (no otherwise). 3: We asked the respondent whether he/she took any loans in 2013-14. Note 4: We elicited characteristics of each field and averaged the responses across fields for each farmer (Note that the sample only includes respondents who own at least one field). 5: Panel C only includes the households who had a soil sample analysis done, 252 farmers.

**Table 1 (cont.): Descriptive statistics of households at baseline in 2014**

| Variable description | N | Mean | St. Dev. |
|---|---|---|---|
| *Panel D - Crop and technology choices* | | | |
| *Crop and seed choices* | | | |
| Cultivated maize in 2013-14 (no=0; yes=1) | 1,000 | 0.96 | 0.19 |
| Cultivate hybrid maize in 2013-14 (no=0; yes=1) | 961 | 0.62 | 0.49 |
| Cultivated groundnut in 2013-14 (no=0; yes=1) | 1,000 | 0.55 | 0.50 |
| Cultivated common bean in 2013-14 (no=0; yes=1) | 1,000 | 0.44 | 0.50 |
| Cultivated soybean in 2013-14 (no=0; yes=1) | 1,000 | 0.46 | 0.50 |
| Inoculated soybean in 2013-14 (no=0; yes=1)[5] | 463 | 0.00 | 0.65 |
| *Land preparation* | | | |
| Used intercropping in 2013-14 (no=0; yes=1) | 1,000 | 0.49 | 0.50 |
| Used crop rotation in the last 5 years (no=0; yes=1) | 1,000 | 0.82 | 0.38 |
| *Fertiliser* | | | |
| Use animal manure in 2013-14 (no=0; yes=1) | 1,000 | 0.29 | 0.45 |
| Incorporated crop residue in 2013-14 (no=0; yes=1) | 1,000 | 0.14 | 0.35 |
| Used inorganic fertiliser in 2013-14 (no=0; yes=1) | 1,000 | 0.88 | 0.31 |
| Planted fertiliser trees in 2013-14 (no=0; yes=1) | 1,000 | 0.09 | 0.28 |
| Used compost in 2013-14 (no=0; yes=1) | 1,000 | 0.019 | 0.13 |
| *Other inputs* | | | |
| Used pesticides in 2013-14 (no=0; yes=1) | 1,000 | 0.00 | 0.00 |
| Used herbicide in 2013-14 (no=0; yes=1) | 1,000 | 0.05 | 0.21 |
| Used fungicide in 2013-14 (no=0; yes=1) | 1,000 | 0.00 | 0.00 |
| *Panel E - Yield expectations* | | | |
| Harvest of maize expected (in kg/ha) | 1,000 | 3,599 | 2,288 |
| Harvest of soy expected (in kg/ha) | 993 | 1,608 | 1,336 |

Notes: Inoculation statistics are conditional cultivating soybean; hybrid maize use is conditional on cultivating maize. Recall that yield expectations were elicited in 50 kg bags per acre, or 50 kg bags per plot. Due to issues with farmers' estimation of acreage, we used the former here, which we converted to kg/ha, excluding outliers above the 95 percentile.

**Table 2: Demonstration plot yields [kg/ha, shelled, adjusted for moisture content], 2014-15 season**

| | N | Mean | St. Dev | Median |
|---|---|---|---|---|
| *Maize* | | | | |
| Maize control | 37 | 2,275 | 1,290 | 2,168 |
| Maize farmer practice | 7 | 2,984 | 958 | 2,806 |
| Maize BPA | 7 | 4,350 | 1,895 | 3,974 |
| Maize BPA + fertilizer trees | 48 | 3,601 | 1,826 | 3,420 |
| *Soybean* | | | | |
| Soybean Control | 22 | 564 | 358 | 543 |
| Soybean BPA | 22 | 995 | 476 | 779 |

There are 19 demonstration plots in 17 villages. Of these 12 are Soybean and Maize plots, 4 are Beans and Maize plots and 3 are Groundnut and Maize plots. The descriptive statistics presented in Table 2 focus on two crops: Soybean and Maize. Appendix A presents the layout of the demonstration plots. One will note that one would have a total of 24 Soybean Control sub-plots and 24 Soybean Best Practice sub-plots in this design. Our sample size is 22 - as one of the demonstration plots did not correctly follow the guidelines. For Maize, which is planted on all three types of demonstration plots, one would have 38 Control sub-plots, 7 Farmer Practice, 7 Best Practice sub-plots, and 48 Best Practice with fertiliser trees sub-plots. Our sample design only has 37 Control sub-plots as, again, one of the demonstration plots did not correctly follow the guidelines.

**Table 3: Descriptive statistics of knowledge questions at endline, in 2015**

| | Question asked | Mean | Std. Dev. |
|---|---|---|---|
| 1 | From the following list, identify which is not a benefit of growing soybeans | 0.60 | 0.49 |
| 2 | True or False: The inoculation of soybean seed enhances nodule formation which in turn enhances plant growth | 0.69 | 0.46 |
| 3 | When mixing inoculant, how many table spoons of sugar should you add to the inoculant bag? | 0.08 | 0.27 |
| 4 | What chemical is best for controlling soybean rust? | 0.16 | 0.36 |
| 5 | When controlling soybean rust, how many millilitre of Folicur should you add to a 15l/16l sprayer? | 0.01 | 0.07 |
| 6 | What chemical is best for controlling insect pests in soya? | 0.36 | 0.48 |
| | | | |
| 7 | From the following list, identify which is not a benefit of growing groundnut. | 0.63 | 0.48 |
| 8 | What is the recommended number of rows per ridges for groundnut? | 0.26 | 0.44 |
| 9 | From the following list, choose the fertiliser used at early flowering stage for groundnut. | 0.10 | 0.30 |
| 10 | From the following list, identify the pesticide which should be used to control for cutworms. | 0.25 | 0.43 |
| 11 | Which of the following options are a sign of groundnut maturity? | 0.98 | 0.13 |
| | | | |
| 12 | From the following options, identify the method which is not used for controlling witch weed | 0.27 | 0.44 |
| 13 | In cm, what is the recommended plant spacing for maize? | 0.09 | 0.28 |
| 14 | Which of the following is not a benefit from covering the field with crop residue, for maize? | 0.54 | 0.50 |
| 15 | How many weeks after planting should you apply urea fertilisers? | 0.41 | 0.49 |
| | | | |
| 16 | Which are not a benefit of soil fertility trees? | 0.73 | 0.45 |
| 17 | True or false: Leaves should be exposed to the sun after the tree has been cut out? | 0.39 | 0.49 |
| 18 | Where exactly on the field should fertiliser trees be planted? | 0.70 | 0.46 |
| 19 | How many weeks after planting the main crops should you plan fertiliser trees? | 0.19 | 0.39 |
| 20 | In which direction should you face when spraying chemicals? | 0.46 | 0.50 |
| | Overall knowledge score | 7.87 | 2.30 |

Note: This table presents the descriptive statistics of the knowledge questions collected one year after the CDI program. The sample includes all 1000 households in 100 villages. All questions are binary yes (1) - no(0) questions. The knowledge tests and answers are presented in the Appendix.

**Table 4: The impact of the CDI program on (planned) adoption of ISFM technologies**

*Linear and Linear Probability Model with Dependent Variables:*

| | Adoption | Soy | Inoculation | Ground-nut | Hybrid Maize | Herbicide | Pesticide | Fungicide | Inorganic fertilizer | Fertiliser tree | Inter-cropping | Animal manure | Crop residue | Compost |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Demonstration plot | 0.793** | 0.054 | 0.166** | 0.098 | 0.162*** | -0.010 | 0.051 | 0.018 | 0.027 | 0.146** | 0.068 | -0.026 | -0.007 | 0.048 |
| | (0.334) | (0.050) | (0.068) | (0.067) | (0.054) | (0.052) | (0.063) | (0.049) | (0.046) | (0.072) | (0.060) | (0.081) | (0.075) | (0.048) |
| Field-day | 0.059 | 0.069 | 0.029 | -0.048 | 0.074 | -0.065 | -0.048 | -0.037 | 0.019 | 0.054 | -0.006 | 0.094 | -0.042 | -0.033 |
| | (0.305) | (0.047) | (0.055) | (0.074) | (0.062) | (0.043) | (0.053) | (0.039) | (0.045) | (0.060) | (0.066) | (0.084) | (0.070) | (0.034) |
| Constant | 5.435*** | 0.875*** | 0.106*** | 0.745*** | 0.780*** | 0.110*** | 0.149*** | 0.082*** | 0.902*** | 0.125*** | 0.804*** | 0.412*** | 0.278*** | 0.067*** |
| | (0.232) | (0.038) | (0.034) | (0.046) | (0.046) | (0.034) | (0.036) | (0.028) | (0.031) | (0.034) | (0.042) | (0.050) | (0.044) | (0.025) |
| Observations | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 |
| R-squared | 0.020 | 0.010 | 0.030 | 0.011 | 0.026 | 0.008 | 0.007 | 0.005 | 0.002 | 0.021 | 0.004 | 0.007 | 0.001 | 0.010 |

Notes: This table present the results of a series of linear regressions with dependent variables: adoption (score out of 13), soy (binary variable), inoculation soy (binary variable), groundnut (binary variable), hybrid maize (binary variable), herbicide (binary variable), pesticide (binary variable), fungicide (binary variable), inorganic fertilizer (binary variable), fertilizer tree (binary variable), intercropping (binary variable), animal manure (binary variable), crop residue (binary variable), and compost (binary variable). These refer to planned adoption in the 2015-16 season. The independent variables are whether or not the individual is in a club which managed a demonstration plot, and whether or not the individual is in a club which was invited to a farmer field day. The estimation uses a two-step procedure. The first steps uses the reported club membership at endline in the treatment villages to predict who would be most likely to join a CDI club. The second step uses all individuals in both treatment and control group whose predicted probability is larger than 0.5. This includes 241 individuals in the treatment group and 173 individuals in the control group. The control variables used in this first step include all village level characteristics included in Appendix Table 1, all household level characteristics included in Appendix Table 4 (and the square terms of the non-binary variables) and baseline adoption indicators. Bootstrapped clustered errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Appendix Table 6 present the results including the control variables in the second step.

**Table 5: The impact of the CDI program on knowledge of ISFM technologies**

*Linear and Linear Probability Model with Dependent Variables:*

| | Knowledge | Knowledge Q1 | Knowledge Q2 | Knowledge Q3 | Knowledge Q4 | Knowledge Q5 | Knowledge Q6 | Knowledge Q7 | Knowledge Q8 | Knowledge Q9 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | |
| Demonstration plot | 0.629 | 0.014 | 0.189*** | -0.003 | 0.107* | -0.004 | 0.102 | 0.059 | 0.036 | -0.008 | |
| | (0.397) | (0.086) | (0.063) | (0.041) | (0.065) | (0.003) | (0.083) | (0.078) | (0.075) | (0.053) | |
| Field-day | -0.140 | 0.018 | 0.070 | -0.018 | 0.058 | 0.019 | -0.038 | -0.082 | -0.065 | -0.060 | |
| | (0.354) | (0.083) | (0.076) | (0.037) | (0.057) | (0.019) | (0.074) | (0.081) | (0.068) | (0.042) | |
| Constant | 7.871*** | 0.600*** | 0.682*** | 0.075*** | 0.122*** | 0.004 | 0.341*** | 0.655*** | 0.278*** | 0.094*** | |
| | (0.227) | (0.054) | (0.047) | (0.024) | (0.032) | (0.003) | (0.047) | (0.049) | (0.046) | (0.035) | |
| Observations | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | |
| R-squared | 0.013 | 0.000 | 0.025 | 0.001 | 0.013 | 0.009 | 0.009 | 0.009 | 0.005 | 0.008 | |
| | Knowledge Q10 | Knowledge Q11 | Knowledge Q12 | Knowledge Q13 | Knowledge Q14 | Knowledge Q15 | Knowledge Q16 | Knowledge Q17 | Knowledge Q18 | Knowledge Q19 | Knowledge Q20 |
| | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) | (19) | (20) | (21) |
| Demonstration plot | 0.138* | -0.010 | 0.036 | 0.085 | -0.181** | -0.029 | -0.001 | 0.025 | 0.049 | 0.051 | -0.025 |
| | (0.075) | (0.017) | (0.078) | (0.057) | (0.084) | (0.087) | (0.072) | (0.081) | (0.074) | (0.066) | (0.083) |
| Field-day | 0.082 | -0.019 | 0.002 | 0.026 | -0.107 | 0.085 | 0.001 | -0.001 | 0.025 | -0.003 | -0.134* |
| | (0.070) | (0.019) | (0.074) | (0.049) | (0.081) | (0.084) | (0.074) | (0.076) | (0.071) | (0.059) | (0.078) |
| Constant | 0.176*** | 0.996*** | 0.278*** | 0.086*** | 0.624*** | 0.443*** | 0.729*** | 0.361*** | 0.694*** | 0.149*** | 0.482*** |
| | (0.040) | (0.006) | (0.046) | (0.030) | (0.051) | (0.054) | (0.045) | (0.049) | (0.046) | (0.039) | -0.048 |
| Observations | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 |
| R-squared | 0.018 | 0.006 | 0.001 | 0.010 | 0.021 | 0.006 | 0.000 | 0.000 | 0.002 | 0.003 | 0.012 |

Notes: This table present the results of a series of linear regressions with dependent variables: knowledge (score out of 20), and the 20 knowledge questions listed in Table 3 (binary variables). The independent variables are whether or not the individual is in a club which managed a demonstration plot, and whether or not the individual is in a club which was invited to a farmer field day. The estimation uses a two-step procedure. The first steps uses the reported club membership at endline in the treatment villages to predict who would be most likely to join a CDI club. The second step uses all individuals in both treatment and control group whose predicted probability is larger than 0.5. This includes 241 individuals in the treatment group and 173 individuals in the control group. The control variables used in this first step include all village level characteristics included in Appendix Table 1, all household level characteristics included in Appendix Table 4 (and the square terms of the non-binary variables) and baseline adoption indicators. Bootstrapped clustered errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Appendix Table 7 present the results including the control variables in the second step.

**Table 6: Correlates of yield expectations for demonstration plot farmers at endline, in 2015**

*OLS Regression of yield expectations at endline [in kg/ha]*

| | Soybean | | Hybrid maize | | Local maize | |
|---|---|---|---|---|---|---|
| Below/above median absolute difference between soil qualities | Similar | Different | Similar | Different | Similar | Different |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Yield expectations at baseline [in kg/ha] | 0.015 | 0.256 | 0.431* | 0.254 | 0.089 | 0.096 |
| | (0.096) | (0.240) | (0.201) | (0.151) | (0.132) | (0.132) |
| Difference between BPA/control subplots on local demonstration plot [in kg/ha] | 0.777** | 0.121 | 1.003*** | 0.292 | 0.093 | 0.218 |
| | (0.292) | (0.404) | (0.155) | (0.454) | (0.204) | (0.378) |
| Observations | 24 | 26 | 40 | 41 | 40 | 41 |
| R-squared | 0.459 | 0.496 | 0.414 | 0.299 | 0.186 | 0.169 |

Notes: This table present the results of a linear regression with dependent variable: Yield expectation at endline. Columns (1) and (2) refer to soybean, Columns (3) and (4) to hybrid maize and Columns (5) and (6) to local maize. We split the sample according to absolute difference in soil quality between the local demonstration plot and the farmer's plot (measured by active carbon in mg/kg): Columns (1), (3) and (5) include the observations below the median of this distribution ("similar"), while Columns (2), (4) and (6) include the observations above the median of this distribution ("different") (the median absolute difference is around 120-130 mg/kg). The yield expectations at baseline refer to the baseline expectations of soy in the case of Columns (1) and (2) and maize in the case of the other Columns. Similarly, the differences between BPA and control subplots refer to the mean differences between these two treatments on the local demonstration plot. Other control variables included but not reported: Gender household head, age household head, education household head (years), number of household members, number of adult household members, maximum education level in the household, acreage of land owned, value of all assets (excluding land) and whether the household cultivated hybrid maize in 2013-15 (for Columns (3) through (6) only). Sample includes the club farmers in the demonstration plot villages. Whether or not farmer is in a club is determined by the self-reported club status at endline. Village-clustered errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

**Table 7: Correlates of knowledge for demonstration plot farmers at endline, in 2015**

*OLS Regression of knowledge score at endline*

| | Soybean | | Hybrid maize | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Yield on BPA subplot on local demonstration plot [in 50 kg bags/acre] | 0.042 | | 0.018** | |
| | (0.028) | | (0.007) | |
| Start of the rain (days) | | -0.018 | | -0.005 |
| | | (0.012) | | (0.009) |
| Total amount of rain (mm) | | 0.001 | | 0.001 |
| | | (0.001) | | (0.001) |
| Number of times flood (flood = More than 50 mm/1 day) | | -0.221* | | -0.152 |
| | | (0.127) | | (0.279) |
| Observations | 61 | 68 | 101 | 101 |
| R-squared | 0.115 | 0.213 | 0.213 | 0.183 |

Notes: This table present the results of a linear regression with dependent variable: Knowledge score at endline. Columns (1) and (2) refer to soybean and Columns (3) and (4) to hybrid maize. The knowledge score for soybean is a number out of 6, while the knowledge score for hybrid maize is a number out of 8. Columns (1) and (3) consider use yield on the BPA subplot as the main independent variable of interest, which refers to the maximum yield on the BPA subplots on the local demonstration plot. Columns (2) and (4) consider various rainfall aggregates as the main independent variables. Other control variables included but not reported: Gender household head, age household head, education household head (years), number of household members, number of adult household members, maximum education level in the household, acreage of land owned, value of all assets (excluding land), relevant yield expectations at baseline, and whether the household cultivated hybrid maize in 2013-15 (for Columns (3) through (6) only). Sample includes the club farmers in the demonstration plot villages. Whether or not farmer is in a club is determined by the self-reported club status at endline. Village-clustered errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

**Table 8: Correlates of knowledge for field-day farmers at endline, in 2015**

*OLS Regression of knowledge score at endline*

| | Soybean | Hybrid maize |
|---|---|---|
| | (1) | (2) |
| Yield expectations at endline [in 50 kg bags/acre] | 0.025** | -0.005 |
| | (0.010) | (0.004) |
| Took credit in 2013-14 season (no=0; yes=1) | 0.337* | 0.134 |
| | (0.186) | (0.203) |
| | | |
| Observations | 77 | 77 |
| R-squared | 0.178 | 0.187 |

Notes: This table present the results of a linear regression with dependent variable: Knowledge score at endline. Column (1) refers to soybean and Column (2) to hybrid maize. The knowledge score for soybean is a number out of 5 (we excluded the first generic question on soybean), while the knowledge score for hybrid maize is a number out of 3 (we focus on the labour-intensive techniques). Other control variables included but not reported: Gender household head, age household head, education household head (years), number of household members, number of adult household members, maximum education level in the household, acreage of land owned, value of all assets (excluding land), relevant yield expectations at baseline, and whether the household cultivated hybrid maize in 2013-15 (for Columns (3) through (6) only). Sample includes the club farmers in the treatment villages (excluding the demonstration plot villages). Whether or not farmer is in a club is determined by the self-reported club status at endline. Village-clustered errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

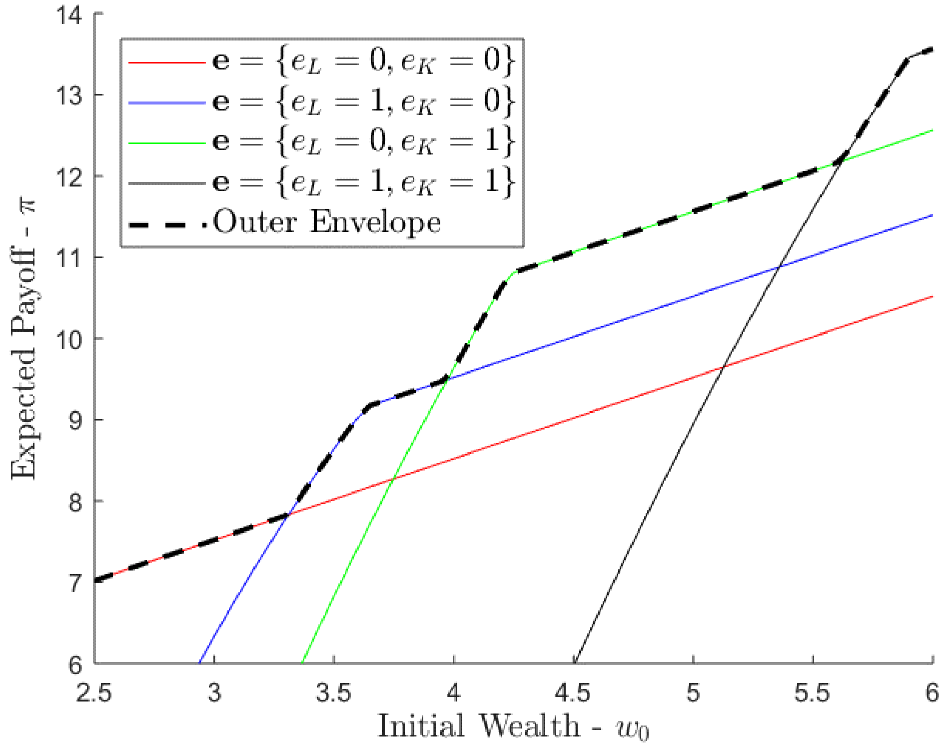



**Figure 1: Predicted relationship between expected payoff and wealth (through applying various levels of effort)**

**Parameters for the model are the following. $\pi_L = 5$; $\pi_K = 5.8$; $\sigma(0) = 3$; $\sigma(1) = 1$; e = 3.**

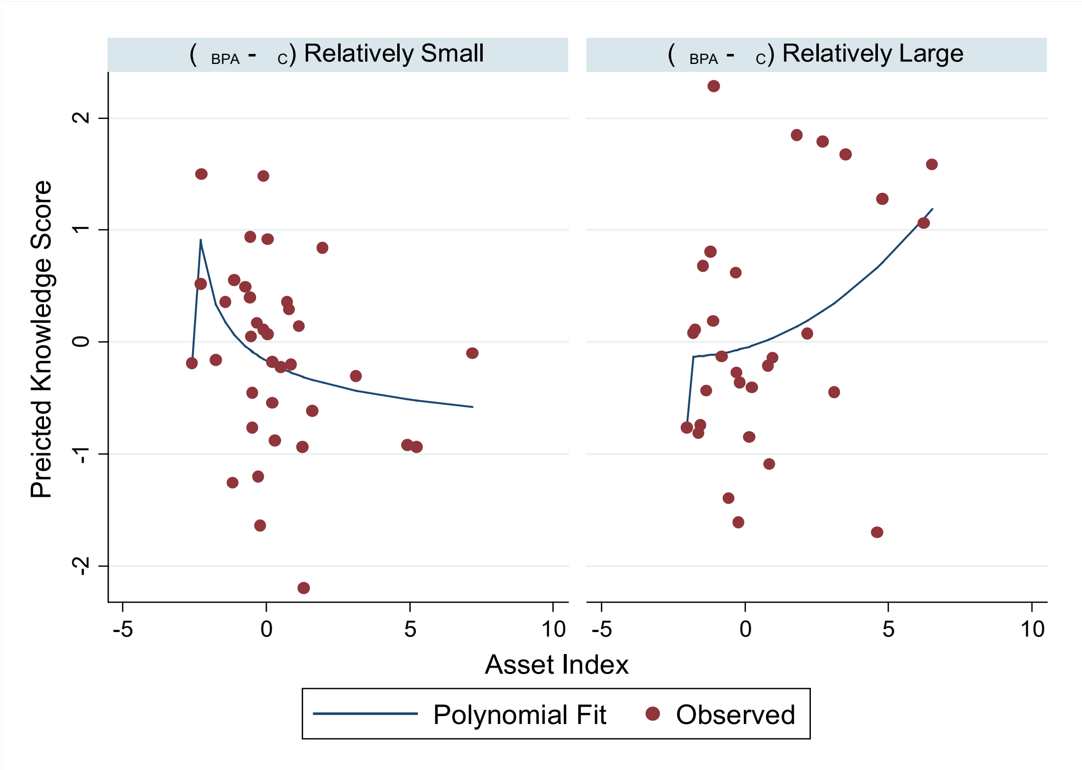



**Figure 2: Capital intensive knowledge increases when priors over profits increase, but only for the wealthy.**

Notes: Includes demonstration plot farmers only.

**Appendix Table 1: Descriptive statistics of villages, balance table, at baseline in 2014**

| | | Full Sample | | | Control villages | | | Treatment villages | | P-value |
|---|---|---|---|---|---|---|---|---|---|---|
| | N | Mean | St. Dev. | N | Mean. | St. Dev | N | Mean. | St. Dev | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Distance to an all-weather/paved road (km) | 100 | 2.22 | 4.40 | 45 | 2.57 | 5.83 | 45 | 1.98 | 2.86 | 0.5436 |
| Distance to a national highway (km) | 100 | 7.83 | 10.45 | 45 | 8.02 | 10.81 | 45 | 8.38 | 11.06 | 0.8763 |
| Distance to a seasonal input market (km) | 100 | 13.26 | 10.72 | 45 | 14.33 | 11.91 | 45 | 12.88 | 10.22 | 0.5381 |
| Distance to an year-round input market (km) | 99 | 5.42 | 5.37 | 45 | 5.42 | 4.53 | 44 | 5.59 | 6.46 | 0.8873 |
| Distance to bank (km) | 98 | 16.71 | 10.97 | 44 | 16.18 | 11.40 | 44 | 17.81 | 11.14 | 0.4972 |
| Distance to seasonal output market (km) | 99 | 4.61 | 6.34 | 45 | 5.26 | 8.29 | 45 | 3.65 | 3.56 | 0.2367 |
| Distance to year-round output market (km) | 100 | 2.49 | 2.81 | 45 | 2.12 | 2.13 | 45 | 2.25 | 2.44 | 0.7888 |
| Number of households | 99 | 61.86 | 44.94 | 45 | 67.56 | 51.38 | 44 | 59.11 | 39.80 | 0.3881 |
| Number of individuals | 95 | 360.31 | 322.09 | 44 | 354.34 | 300.54 | 41 | 395.78 | 370.17 | 0.5743 |
| Percentage of women | 90 | 58.27 | 10.65 | 43 | 56.91 | 12.64 | 37 | 60.16 | 8.61 | 0.1776 |
| Size of the largest ethnic group (in percentage) | 100 | 91.16 | 13.38 | 45 | 86.8* | 14.73 | 45 | 94.47* | 11.21 | 0.0068 |
| Number of ethnic groups | 100 | 1.24 | 0.57 | 45 | 1.31 | 0.60 | 45 | 1.16 | 0.52 | 0.1908 |
| Area under soy cultivation (percentage) | 100 | 24.73 | 18.31 | 45 | 25.84 | 21.05 | 45 | 22.89 | 14.13 | 0.4366 |
| Farmers cultivating soy (percentage) | 99 | 45.06 | 28.76 | 44 | 45.23 | 30.30 | 45 | 43.47 | 28.93 | 0.7799 |
| Number of visits by govt. extension agents (last year) | 100 | 7.68 | 17.75 | 45 | 9.31 | 21.09 | 45 | 7.04 | 15.94 | 0.5668 |
| Number of visits by NGO extension agents (last year) | 97 | 15.65 | 24.57 | 45 | 17.91 | 31.78 | 42 | 13.74 | 16.31 | 0.4393 |
| Number of organisations active in the village | 100 | 1.71 | 1.36 | 45 | 1.56 | 1.42 | 45 | 1.82 | 1.28 | 0.3533 |
| Average number of members per organisation | 100 | 10.78 | 11.42 | 45 | 8.39* | 9.89 | 45 | 12.97* | 11.87 | 0.0502 |
| Daily wage rate at planting time | 100 | 1326.07 | 1465.90 | 45 | 1130.00 | 1296.43 | 45 | 1531.76 | 1685.97 | 0.2086 |
| Daily wage rate outside of peak times | 100 | 985.71 | 1351.73 | 45 | 703.8* | 580.04 | 45 | 1264.44* | 1885.79 | 0.0621 |
| Daily wage rate at harvesting time | 100 | 1033.00 | 1000.06 | 45 | 928.89 | 938.90 | 45 | 1200.00 | 1134.13 | 0.2201 |
| Distance to EPA capital | 100 | 15.73 | 5.54 | 45 | 16.11 | 5.28 | 45 | 15.39 | 5.96 | 0.5443 |

This table presents the descriptive statistics of the village questionnaire for the 100 villages revisited one year after baseline (Columns 1 through 3), the control villages (Columns 4 through 6) and the treatment villages (Columns 7 through 9). Column (10) presents the P-value of a t-test with unequal variances between the treatment and control group. Note that Columns (4) through (10) exclude the purposefully selected sample of 10 demonstration plot villages and present comparisons for the randomly selected sample of 90 villages only. The t-tests give similar results if one includes all 250 baseline villages.

**Appendix Table 2: Descriptive statistics of villages, by demonstration plot and CDI club membership status, at baseline in 2014**

| | Demo plot villages | | | Non-demo plot villages | | | P-value |
|---|---|---|---|---|---|---|---|
| | N | Mean. | St. Dev | N | Mean. | St. Dev | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Distance to an all-weather/paved road (km) | 17 | 1.95 | 2.57 | 39 | 1.87 | 2.86 | 0.9196 |
| Distance to a national highway (km) | 17 | 5.68 | 6.17 | 39 | 8.61 | 11.41 | 0.2216 |
| Distance to a seasonal input market (km) | 17 | 12.38 | 10.09 | 39 | 13.08 | 10.48 | 0.8149 |
| Distance to an year-round input market (km) | 17 | 5.48 | 4.99 | 38 | 5.79 | 6.84 | 0.8498 |
| Distance to bank (km) | 17 | 14.21 | 8.66 | 38 | 18.53 | 11.21 | 0.1273 |
| Distance to seasonal output market (km) | 16 | 7.13 | 10.12 | 39 | 3.75 | 3.60 | 0.2098 |
| Distance to year-round output market (km) | 17 | 3.47 | 4.45 | 39 | 2.45 | 2.54 | 0.3885 |
| Number of households | 17 | 50.82 | 31.21 | 38 | 59.08 | 41.50 | 0.4199 |
| Number of individuals | 16 | 330.69 | 191.24 | 36 | 379.08 | 388.94 | 0.5507 |
| Percentage of women | 16 | 57.61 | 8.07 | 32 | 60.81 | 8.51 | 0.2119 |
| Size of the largest ethnic group (in percentage) | 17 | 94.65 | 11.34 | 39 | 94.90 | 10.99 | 0.9394 |
| Number of ethnic groups | 17 | 1.29 | 0.69 | 39 | 1.13 | 0.47 | 0.3729 |
| Area under soy cultivation (percentage) | 17 | 26.94 | 20.98 | 39 | 23.38 | 14.23 | 0.53 |
| Farmers cultivating soy (percentage) | 17 | 53.24 | 26.45 | 39 | 42.21 | 28.16 | 0.1691 |
| Number of visits by govt. extension agents (last year) | 17 | 4.29 | 5.58 | 39 | 7.15 | 16.86 | 0.348 |
| Number of visits by NGO extension agents (last year) | 16 | 11.19 | 13.47 | 37 | 14.54 | 16.93 | 0.4479 |
| Number of organisations active in the village | 17 | 1.82 | 1.29 | 39 | 1.82 | 1.32 | 0.9937 |
| Average number of members per organisation | 17 | 12.28 | 16.40 | 39 | 12.70 | 10.14 | 0.9237 |
| Daily wage rate at planting time | 17 | 1284.00 | 1287.73 | 39 | 1664.85 | 1773.45 | 0.3722 |
| Daily wage rate outside of peak times | 17 | 811.76 | 484.62 | 39 | 1387.18 | 1999.70 | 0.0983 |
| Daily wage rate at harvesting time | 17 | 817.65 | 433.35 | 39 | 1246.15 | 1193.98 | 0.0548 |
| Distance to EPA capital | 17 | 16.38 | 4.88 | 39 | 15.05 | 6.08 | 0.3936 |

This table presents the descriptive statistics of the village questionnaire for all demonstration plot villages (Columns (1) through (3)) and the complementary set of non-demonstration plot villages (Columns (4) through (6)). Column (7) presents the P-value of a t-test with unequal variances between these two groups. The rest of the table includes the villages which formed CDI clubs (Columns (8) through (10)) and the complementary set of villages which did not form CDI clubs (Columns (11) through (13)). The sample includes all villages which were revisited after one year and belong to the treatment group, N=55.

**Appendix Table 2: Descriptive statistics of villages, by demonstration plot and CDI club membership status, at baseline in 2014 (cont.)**

| | CDI club villages | | | Non-CDI club villages | | | P-value |
|---|---|---|---|---|---|---|---|
| | N | Mean. | St. Dev | N | Mean. | St. Dev | |
| | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Distance to an all-weather/paved road (km) | 48 | 1.91 | 2.90 | 8 | 1.81 | 1.76 | 0.8959 |
| Distance to a national highway (km) | 48 | 8.68 | 10.66 | 8 | 1.98 | 1.21 | 0.0001 |
| Distance to a seasonal input market (km) | 48 | 13.76 | 10.69 | 8 | 7.53 | 4.92 | 0.0142 |
| Distance to an year-round input market (km) | 47 | 5.74 | 6.67 | 8 | 5.40 | 3.51 | 0.8307 |
| Distance to bank (km) | 47 | 17.52 | 10.71 | 8 | 15.31 | 10.44 | 0.5948 |
| Distance to seasonal output market (km) | 47 | 4.56 | 6.48 | 8 | 5.76 | 5.56 | 0.5915 |
| Distance to year-round output market (km) | 48 | 2.81 | 3.41 | 8 | 2.50 | 1.91 | 0.7186 |
| Number of households | 47 | 57.60 | 39.32 | 8 | 50.25 | 35.07 | 0.6023 |
| Number of individuals | 45 | 378.49 | 359.17 | 7 | 272.29 | 144.08 | 0.1792 |
| Percentage of women | 40 | 59.12 | 7.71 | 8 | 62.88 | 11.48 | 0.4003 |
| Size of the largest ethnic group (in percentage) | 48 | 94.06 | 11.70 | 8 | 99.38 | 1.77 | 0.0047 |
| Number of ethnic groups | 48 | 1.21 | 0.58 | 8 | 1.00 | 0.00 | 0.0168 |
| Area under soy cultivation (percentage) | 48 | 25.44 | 17.19 | 8 | 18.63 | 9.93 | 0.1336 |
| Farmers cultivating soy (percentage) | 48 | 46.81 | 27.59 | 8 | 38.00 | 30.39 | 0.4615 |
| Number of visits by govt. extension agents (last year) | 48 | 5.81 | 14.68 | 8 | 9.13 | 13.01 | 0.5275 |
| Number of visits by NGO extension agents (last year) | 45 | 12.40 | 15.85 | 8 | 19.88 | 15.72 | 0.2448 |
| Number of organisations active in the village | 48 | 1.98 | 1.31 | 8 | 0.88 | 0.64 | 0.0014 |
| Average number of members per organisation | 48 | 12.46 | 11.65 | 8 | 13.25 | 16.18 | 0.898 |
| Daily wage rate at planting time | 48 | 1613.08 | 1746.25 | 8 | 1166.13 | 662.60 | 0.2049 |
| Daily wage rate outside of peak times | 48 | 1145.83 | 1580.26 | 8 | 1612.50 | 2413.32 | 0.6115 |
| Daily wage rate at harvesting time | 48 | 1129.17 | 1089.74 | 8 | 1037.50 | 704.96 | 0.7606 |
| Distance to EPA capital | 48 | 15.86 | 5.93 | 8 | 13.04 | 3.80 | 0.0999 |

This table presents the descriptive statistics of the village questionnaire for all demonstration plot villages (Columns (1) through (3)) and the complementary set of non-demonstration plot villages (Columns (4) through (6)). Column (7) presents the P-value of a t-test with unequal variances between these two groups. The rest of the table includes the villages which formed CDI clubs (Columns (8) through (10)) and the complementary set of villages which did not form CDI clubs (Columns (11) through (13)). The sample includes all villages which were revisited

**Appendix Table 3: Descriptive statistics of treatment households, by treatment group, at baseline in 2014**

| Variable | Treatment villages | | | Villages without clubs | | |
|---|---|---|---|---|---|---|
| | Obs. | Mean | St. Dev. | Obs. | Mean | St. Dev. |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Gender of household head (0=male; 1=female) | 450 | 0.18 | 0.38 | 120 | 0.22 | 0.41 |
| Age of household head (years) | 450 | 41.79 | 15.22 | 120 | 41.84 | 13.99 |
| Education of household head (years of education) | 450 | 4.52 | 3.40 | 120 | 4.34 | 3.27 |
| Number of household members | 450 | 5.19 | 2.18 | 120 | 5.28 | 1.97 |
| Number of adult household members | 450 | 2.28 | 0.93 | 120 | 2.23 | 0.90 |
| Max years of education in household (years) | 450 | 6.97 | 2.94 | 120 | 6.88 | 2.73 |
| Total owned land (acres) | 450 | 5.17 | 7.68 | 120 | 3.98 | 4.10 |
| Total value of all assets (excluding land) | 450 | 147816 | 478190 | 120 | 83405 | 190480 |
| Are government extension agents one of three main sources of information (no=0; yes=1) | 450 | 0.32 | 0.47 | 120 | 0.19 | 0.40 |
| Are other farmers one of three main sources of information (no=0; yes=1) | 450 | 0.75 | 0.43 | 120 | 0.73 | 0.45 |
| Took credit in 2013-14 season(no=0; yes=1) | 450 | 0.18 | 0.38 | 120 | 0.10 | 0.30 |
| Ability to take credit in 2014-15 season (no=0; yes=1) | 450 | 0.47 | 0.50 | 120 | 0.47 | 0.50 |

This table presents the descriptive statistics of the households in the treatment midline villages (Columns 1 through 3 - which is identical to the Columns (7)-(9) in Table 3), for villages without clubs (Columns 4 through 6), the non-club households in the villages with clubs (Columns 7 through 9) and the club households in the villages with clubs (Columns (10) through (12)). Column (13) presents the P-value of a t-test with unequal variances between the first and second sub-group; Column (14) presents the P-value of a t-test with unequal variances between the second and third sub-group. Note that this table only includes the households in the villages in the random sample, and not in the purposefully selected sample of 10 demonstration plot villages. Total sample size = 450 households in 45 villages.

**Appendix Table 3: Descriptive statistics of treatment households, by treatment group, at baseline in 2014 (cont.)**

| Variable | Villages with clubs Non-club members | | | Villages with clubs Club members | | | | |
|---|---|---|---|---|---|---|---|---|
| | Obs. | Mean | St. Dev. | Obs. | Mean | St. Dev. | | |
| | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Gender of household head (0=male; 1=female) | 164 | 0.15 | 0.36 | 166 | 0.17 | 0.38 | 0.174 | 0.689 |
| Age of household head (years) | 164 | 42.96 | 17.35 | 166 | 40.60 | 13.76 | 0.5494 | 0.172 |
| Education of household head (years of education) | 164 | 4.13 | 3.46 | 166 | 5.04 | 3.40 | 0.6066 | 0.0171 |
| Number of household members | 164 | 4.80 | 2.21 | 166 | 5.51 | 2.25 | 0.0573 | 0.0042 |
| Number of adult household members | 164 | 2.24 | 0.84 | 166 | 2.34 | 1.04 | 0.9197 | 0.3679 |
| Max years of education in household (years) | 164 | 6.61 | 3.14 | 166 | 7.38 | 2.86 | 0.4345 | 0.02 |
| Total owned land (acres) | 164 | 4.48 | 5.12 | 166 | 6.72 | 10.89 | 0.3615 | 0.0172 |
| Total value of all assets (excluding land) | 164 | 169816 | 578586 | 166 | 172645 | 511133 | 0.0757 | 0.9625 |
| Are government extension agents one of three main sources of information (no=0; yes=1) | 164 | 0.34 | 0.48 | 166 | 0.40 | 0.49 | 0.0041 | 0.2923 |
| Are other farmers one of three main sources of information (no=0; yes=1) | 164 | 0.73 | 0.45 | 166 | 0.80 | 0.40 | 0.991 | 0.1397 |
| Took credit in 2013-14 season(no=0; yes=1) | 164 | 0.17 | 0.38 | 166 | 0.25 | 0.43 | 0.0804 | 0.0888 |
| Ability to take credit in 2014-15 season (no=0; yes=1) | 164 | 0.45 | 0.50 | 166 | 0.49 | 0.50 | 0.7973 | 0.5053 |

This table presents the descriptive statistics of the households in the treatment midline villages (Columns 1 through 3 - which is identical to the Columns (7)-(9) in Table 3), for villages without clubs (Columns 4 through 6), the non-club households in the villages with clubs (Columns 7 through 9) and the club households in the villages with clubs (Columns (10) through (12)). Column (13) presents the P-value of a t-test with unequal variances between the first and second sub-group; Column (14) presents the P-value of a t-test with unequal variances between the second and third sub-group. Note that this table only includes the households in the villages in the random sample, and not in the purposefully selected sample of 10 demonstration plot villages. Total sample size = 450 households in 45 villages.

**Appendix Table 4: Descriptive statistics of farmer club members, by activity, at baseline in 2014**

| | Do not participate in fieldday | | | Participate in fieldday | | | |
|---|---|---|---|---|---|---|---|
| Variable | Obs. | Mean | St. Dev. | Obs. | Mean | St. Dev. | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Gender of household head (0=male; 1=female) | 147 | 0.18 | 0.38 | 53 | 0.09 | 0.30 | 0.11 |
| Age of household head (years) | 147 | 42.35 | 14.82 | 53 | 42.15 | 11.84 | 0.9234 |
| Education of household head (years of education) | 147 | 4.38*** | 3.29 | 53 | 6.32*** | 3.64 | 0.001 |
| Number of household members | 147 | 5.35 | 2.34 | 53 | 5.77 | 1.90 | 0.198 |
| Number of adult household members | 147 | 2.35 | 1.03 | 53 | 2.42 | 0.89 | 0.6799 |
| Max years of education in household (years) | 147 | 6.75*** | 3.17 | 53 | 8.70*** | 2.18 | 0 |
| Total owned land (acres) | 147 | 5.45** | 9.23 | 53 | 9.61** | 11.46 | 0.0197 |
| Total value of all assets (excluding land) | 147 | 101376.70 | 149235.90 | 53 | 352206.70 | 859917.50 | 0.0394 |
| Are government extension agents one of three main sources of information (no=0; yes=1) | 147 | 0.39 | 0.49 | 53 | 0.51 | 0.50 | 0.1558 |
| Are other farmers one of three main sources of information (no=0; yes=1) | 147 | 0.78 | 0.42 | 53 | 0.85 | 0.36 | 0.2266 |
| Took credit in 2013-14 season(no=0; yes=1) | 147 | 0.22*** | 0.41 | 53 | 0.43*** | 0.50 | 0.0061 |
| Ability to take credit in 2014-15 season (no=0; yes=1) | 147 | 0.46* | 0.50 | 53 | 0.60* | 0.49 | 0.0654 |

This table presents the descriptive statistics of the households in the CDI clubs. In Columns (1) through (7), we distinguish between households who participated in fielddays and households who do not participate in fielddays. Column (7) presents the P-value of a t-test with unequal variances between these two groups. In Columns (2) through (14) we distinguish between households who participate in demo plots and households who do not participate in demo plots. Column (14) presents the P-value of a t-test with unequal variances between these two groups. The full sample corresponds to all self-declared club members in the treatment villages (including Tongolele) in the 100 vilages.

**Appendix Table 4: Descriptive statistics of farmer club members, by activity, at baseline in 2014 (cont.)**

| | Do not participate in demo plot | | | Participate in demo plot | | | |
|---|---|---|---|---|---|---|---|
| Variable | Obs. | Mean | St. Dev. | Obs. | Mean | St. Dev. | |
| | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Gender of household head (0=male; 1=female) | 134 | 0.16 | 0.36 | 66 | 0.15 | 0.36 | 0.9241 |
| Age of household head (years) | 134 | 41.37 | 14.50 | 66 | 44.17 | 13.05 | 0.1723 |
| Education of household head (years of education) | 134 | 4.73 | 3.44 | 66 | 5.24 | 3.57 | 0.3342 |
| Number of household members | 134 | 5.57 | 2.38 | 66 | 5.26 | 1.90 | 0.3216 |
| Number of adult household members | 134 | 2.37 | 1.08 | 66 | 2.36 | 0.78 | 0.9434 |
| Max years of education in household (years) | 134 | 7.28 | 2.91 | 66 | 7.25 | 3.38 | 0.9511 |
| Total owned land (acres) | 134 | 6.88 | 11.74 | 66 | 5.90 | 4.92 | 0.4049 |
| Total value of all assets (excluding land) | 134 | 162225.10 | 521748.10 | 66 | 179260.10 | 349419.00 | 0.7848 |
| Are government extension agents one of three main sources of information (no=0; yes=1) | 134 | 0.36*** | 0.48 | 66 | 0.56*** | 0.50 | 0.0074 |
| Are other farmers one of three main sources of information (no=0; yes=1) | 134 | 0.79 | 0.41 | 66 | 0.80 | 0.40 | 0.8436 |
| Took credit in 2013-14 season(no=0; yes=1) | 134 | 0.22** | 0.41 | 66 | 0.39** | 0.49 | 0.013 |
| Ability to take credit in 2014-15 season (no=0; yes=1) | 134 | 0.46 | 0.50 | 66 | 0.56 | 0.50 | 0.1953 |

This table presents the descriptive statistics of the households in the CDI clubs. In Columns (1) through (7), we distinguish between households who participated in fielddays and households who do not participate in fielddays. Column (7) presents the P-value of a t-test with unequal variances between these two groups. In Columns (2) through (14) we distinguish between households who participate in demo plots and households who do not participate in demo plots. Column (14) presents the P-value of a t-test with unequal variances between these two groups. The full sample corresponds to all self-declared club members in the treatment villages (including Tongolele) in the 100 vilages.

**Appendix Table 5: The impact of the CDI program on (planned) adoption of ISFM technologies (with additional control variables)**

*Linear and Linear Probability Model with Dependent Variables:*

| | Adoption | Soy | Inoculation | Ground-nut | Hybrid Maize | Herbicide | Pesticide | Fungicide | Inorganic fertilizer | Fertiliser tree | Inter-cropping | Animal manure | Crop residue | Compost |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| Demonstration plot | 0.810** | 0.017 | 0.139* | 0.101 | 0.101 | 0.017 | 0.077 | 0.050 | 0.052 | 0.191** | 0.022 | 0.061 | -0.024 | 0.006 |
| | (0.385) | (0.065) | (0.071) | (0.084) | (0.069) | (0.059) | (0.084) | (0.053) | (0.054) | (0.088) | (0.078) | (0.107) | (0.089) | (0.054) |
| Field-day | -0.054 | 0.058 | -0.006 | -0.065 | 0.041 | -0.031 | -0.063 | -0.019 | 0.040 | 0.053 | -0.004 | 0.059 | -0.042 | -0.075 |
| | (0.381) | (0.059) | (0.065) | (0.097) | (0.081) | (0.055) | (0.084) | (0.048) | (0.066) | (0.077) | (0.084) | (0.112) | (0.094) | (0.052) |
| Controls included | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 |
| R-squared | 0.218 | 0.238 | 0.194 | 0.217 | 0.198 | 0.157 | 0.122 | 0.124 | 0.135 | 0.169 | 0.226 | 0.21 | 0.229 | 0.213 |

Notes: This table present the results of a series of linear regressions with dependent variables: adoption (score out of 13), soy (binary variable), inoculation soy (binary variable), groundnut (binary variable), hybrid maize (binary variable), herbicide (binary variable), pesticide (binary variable), fungicide (binary variable), inorganic fertilizer (binary variable), fertilizer tree (binary variable), intercropping (binary variable), animal manure (binary variable), crop residue (binary variable), and compost (binary variable). These refer to planned adoption in the 2015-16 season. The independent variables are whether or not the individual is in a club which managed a demonstration plot, and whether or not the individual is in a club which was invited to a farmer field day. The estimation uses a two-step procedure. The first steps uses the reported club membership at endline in the treatment villages to predict who would be most likely to join a CDI club. The second step uses all individuals in both treatment and control group whose predicted probability is larger than 0.5. This includes 241 individuals in the treatment group and 173 individuals in the control group. The control variables used in both steps include all village level characteristics included in Appendix Table 1, all household level characteristics included in Appendix Table 3 (and the square terms of the non-binary variables) and baseline adoption indicators. Bootstrapped clustered errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

**Appendix Table 6: The impact of the CDI program on knowledge of ISFM technologies (with additional control variables)**

Linear and Linear Probability Model with Dependent Variables:

| | Knowledge | Knowledge Q1 | Knowledge Q2 | Knowledge Q3 | Knowledge Q4 | Knowledge Q5 | Knowledge Q6 | Knowledge Q7 | Knowledge Q8 | Knowledge Q9 |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Demonstration plot | 0.115 | -0.013 | 0.169* | -0.008 | 0.123 | -0.007 | 0.039 | 0.047 | -0.018 | -0.057 |
| | (0.477) | (0.098) | (0.097) | (0.056) | (0.078) | (0.011) | (0.103) | (0.103) | (0.096) | (0.067) |
| Field day | -0.480 | 0.047 | 0.050 | 0.008 | 0.032 | 0.015 | -0.060 | -0.076 | -0.112 | -0.091* |
| | (0.489) | (0.120) | (0.089) | (0.054) | (0.081) | (0.023) | (0.100) | (0.114) | (0.090) | (0.056) |
| Controls included | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 |
| R-squared | 0.189 | 0.120 | 0.142 | 0.094 | 0.112 | 0.124 | 0.179 | 0.134 | 0.140 | 0.160 |

| | Knowledge Q10 | Knowledge Q11 | Knowledge Q12 | Knowledge Q13 | Knowledge Q14 | Knowledge Q15 | Knowledge Q16 | Knowledge Q17 | Knowledge Q18 | Knowledge Q19 | Knowledge Q20 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) | (19) | (20) | (21) |
| Demonstration plot | 0.057 | -0.003 | -0.041 | 0.070 | -0.186* | -0.016 | -0.016 | 0.008 | 0.014 | -0.006 | -0.041 |
| | (0.098) | (0.025) | (0.094) | (0.068) | (0.110) | (0.105) | (0.087) | (0.108) | (0.095) | (0.084) | (0.107) |
| Field-day | 0.022 | -0.025 | -0.050 | -0.040 | -0.073 | 0.075 | -0.003 | -0.027 | -0.008 | -0.047 | -0.117 |
| | (0.093) | (0.031) | (0.096) | (0.054) | (0.119) | (0.111) | (0.100) | (0.111) | (0.095) | (0.078) | (0.110) |
| Controls included | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 | 414 |
| R-squared | 0.182 | 0.119 | 0.158 | 0.224 | 0.164 | 0.158 | 0.169 | 0.089 | 0.141 | 0.129 | 0.128 |

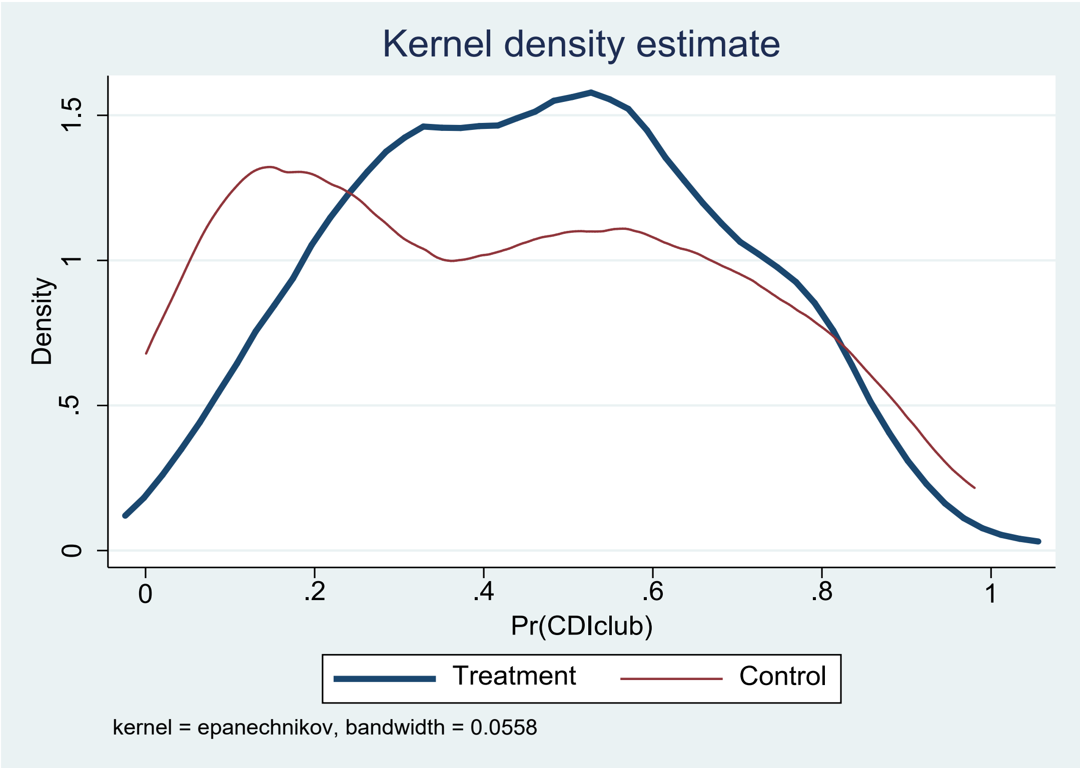
Notes: This table present the results of a series of linear regressions with dependent variables: knowledge (score out of 20), and the 20 knowledge questions listed in Table 3 (binary variables). The independent variables are whether or not the individual is in a club which managed a demonstration plot, and whether or not the individual is in a club which was invited to a farmer field day. The estimation uses a two-step procedure. The first steps uses the reported club membership at endline in the treatment villages to predict who would be most likely to join a CDI club. The second step uses all individuals in both treatment and control group whose predicted probability is larger than 0.5. This includes 241 individuals in the treatment group and 173 individuals in the control group. The control variables used in both steps include all village level characteristics included in Appendix Table 1, all household level characteristics included in Appendix Table 3 (and the square terms of the non-binary variables) and baseline adoption indicators. Bootstrapped clustered errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

**Appendix Table 7: The impact of the CDI program on (planned) adoption of ISFM technologies - Farmer Fixed Effects Approach**

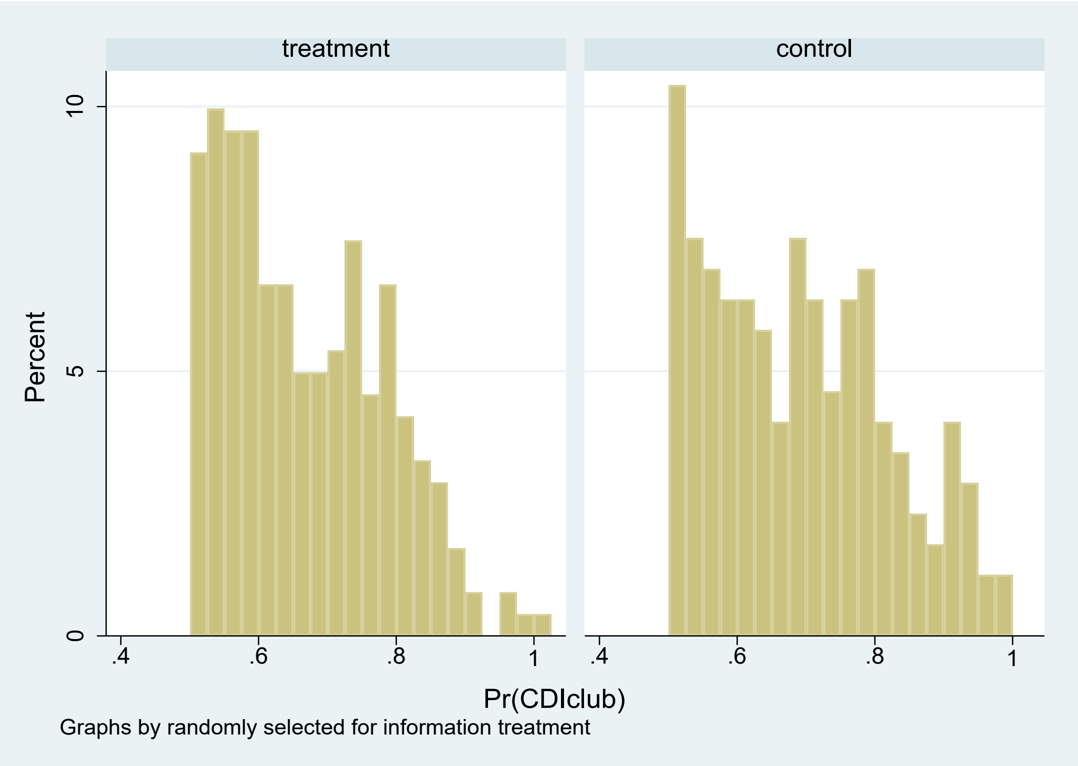
*Panel Estimation: Dependent Variable: Adoption Score (out of 13)*

| | (1) |
|---|---|
| Demonstration plot | 0.336 |
| | (0.227) |
| Field-day | -0.074 |
| | (0.231) |
| 2015-16 Year | 1.911*** |
| | (0.111) |
| Constant | 3.599*** |
| | (0.04) |
| | |
| Observations | 1,898 |
| Number of households | 949 |
| R-squared | 0.432 |

Notes: This table present the results of a farmer fixed-effects regression with dependent variable (xtreg, fe): Adoption score (score out of 13). The independent variables are whether or not the individual is in a club which managed a demonstration plot, and whether or not the individual is in a club which was invited to a farmer field day. We include, but do not report, the farmers who live in treatment villages but do not belong to the treatment clubs in this regression. We use adoption of 2013-14 as elicited at baseline (in 2014), and planned adoption elicited in 2005 referring to the 2015-16 season. Standard errors clustered at the village level in parentheses. Whether or not farmer is in a club is determined by the self-reported club status in 2015.*** p<0.01, ** p<0.05, * p<0.1

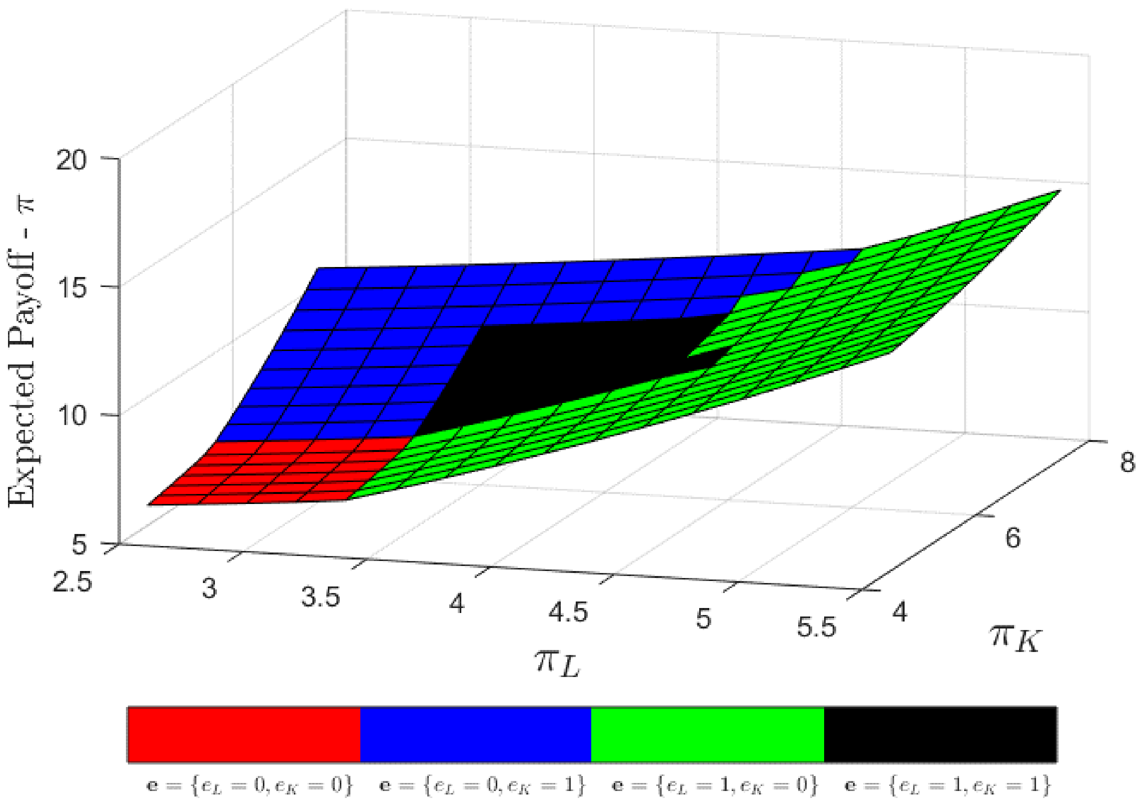



**Appendix Figure 1: Kernel density plot of the predicted probability of CDI club membership in the treatment and control villages (see also Table 4)**




**Appendix Figure 2: The distribution of the probability of CDI club membership, by treatment and control villages, for the sample for which the probability estimated is larger than 0.5 (see also Table 4)**


**Appendix Figure 3: Rainfall distributions in the demonstration plot villages**




**Appendix Figure 4: Predicted relationship between expected payoff and profits from capital and labor intensive technology (through applying various levels of effort)**

**Parameters for the model are the following. $w_0$ = 5; σ(0) = 3; σ(1) = 1; e = 3.**

**Appendix A: Demonstration ploy lay-out**

| **Bean and Maize Demonstration Plot Layout -2014/2015 Season** | | **Soybean-ISFM Demonstration Plot Layout-2014/2015 Season** | | **Groundnut and Maize Demonstration Plot Layout -2014/2015 Season** | |
|---|---|---|---|---|---|
| **Block 1** | **Block 2** | **Block 1** | **Block 2** | **Block 1** | **Block 2** |
| Bean Farmer Practice | Maize Farmer Practice | T2: Hybrid Maize Control | T1:Hybrid Maize/fertilizer trees/ fertilizer /CA | Groundnut Farmer Practice | Maize Farmer Practice |
| Bean Best Practice Agronomy | Maize Best Practice Agronomy | T4: Improved Soybean control | T5: Soybean Best Practice Agronomy | Groundnut Best Practice Agronomy | Maize Best Practice Agronomy |
| Maize Control | Maize Control | T1:Hybrid Maize/fertilizer trees / fertilizer /CA | T2: Hybrid Maize Control | Maize Control | Maize Control |
| Beans Control | Beans Control | T5: Soybean Best Practice Agronomy | T3: Crop rotation with T5 (Hybrid Maize/fertilizer trees /fertilizer/CA) | Groundnut Control | Groundnut Control |
| | | T3: Crop rotation with T5 (Hybrid Maize/ fertilizer trees (fertilizer/CA) | T4: Improved Soybean control | | |

## Appendix B

### KNOWLEDGE MODULE

**<u>Enumerator instructions:</u>** State the following: The following questions will ask you to provide answers associated with various farming practices. We are not here to judge how much you know or don't know about farming a particular crop. We are only trying to understand how much information is shared with you. If you are unsure of how you would respond to any of the next questions, feel free to say 'I don't know.'"

**Enumerator instructions**: Read the 20 questions below out loud and write down the answer of the respondent.

-99 = DON'T KNOW

**<u>SOYBEANS</u>**

D1 From the following list, identify which is not a benefit of growing soy beans:

1 = CAN BE USED TO PRODUCE COOKING OIL
2 = IMPROVES SOIL FERTILITY
**3 = DOES NOT REQUIRE THE USE OF HERBICIDES AND PESTICIDES**
4 = CAN BE USED IN THE PREPARATION OF A VARIETY OF FOODS

D2 ***True or false*** : The inoculation of soybean seed enhances nodule formation which in turn enhances plant growth **Answer: True**

1 = TRUE
2 = FALSE

D3 When mixing inoculant, how many table spoons of sugar should you add to the inoculant bag?
**Answer: 2**

D4 What chemical is best for controlling soya rust?
**Answer: Folicur, 1**

1 = FOLICUR
2 = CYPERMETHRIN
3 = HARNESS
4 = ROUND-UP

D5 When controlling soy bean rust, how many millilitres of Folicur should you add to a 15/16lL sprayer?
**Answer: 25 ml**

D6 What chemical is best for controlling pests in soya?
**Answer: Cypermethylin, 1**

1 = CYPERMETHYLIN
2 = KARATE
3 = FOLICUR

**<u>GROUNDNUT</u>**

D7 From the following list, identify which is not a benefit of growing groundnut:

1 = PRODUCES COOKING OIL

2 = PRODUCES FEED FOR LIVESTOCK
3= CAN BE USED IN THE PREPARATION OF A VARIETY OF FOODS
**4 = GROUNDNUT IS FLOOD RESISTANT (LIKE RICE)**

D8 What is the recommended number of rows per ridge under best practice agronomy for groundnuts? -99 = DON'T KNOW

**Answer: 2 rows**

D9 From the following list, choose the fertiliser used at the early flowering stage in groundnut production :

**Answer: Gypsum, 3**

1 = SINGLE SUPERPHOSPATE
2 = D COMPOUND
3 = GYPSUM

D10 From the following list, identify the pesticide which should be used to control for cutworms:

**Answer: Karate, 1**

1 = KARATE
2 = MONOCRON
3 = DIAMTHOATE

D11 Which of the following options are a sign of groundnut maturity:

1 = **THE LEAVES TURN YELLOW AND BEGIN TO FALL OFF**
2 = THE GROUNDNUT SHELL BECOMES VERY SOFT
3 **= THE NUTS BECOME SMOOTH**

**MAIZE**

D12 From the following options, identify the method that is not used for controlling witch weed?

1 = MULCHING
2 = CROP ROTATION
3 = MANURE APPLICATION
**4 = APPLICATION OF FOLICUR**

D13 In centimetres, what is the recommended plant spacing for Maize under Best practice agronomy?

**Answer: 25 cm**

D14 Which of the following is not a benefit of planting maize in fields covered with crop residues (conservation agriculture):

1= IMPROVED SOIL STRUCTURE
**2 = REDUCES SOIL FERTILITY**
3 = IMPROVED SOIL WATER RETENTION
4 = HELP TO CONTROL WEEDS

D15 How many weeks after planting should you apply UREA fertilizer?

**Answer: 3 weeks**

**SOIL FERTILITY ENHANCING TREES**

D16 Which are not benefits of these trees: -99 = DON'T

KNOW

1 = LEAVES INCREASE SOIL FERTILITY
2 = ROOTS HELP IMPROVE SOIL STRUCTURE
**3 = GOATS LIKE TO EAT THEM**

D17 True or false: Leaves should be exposed to the sun after the tree has been cut

**Answer: true**

D18 Where exactly on the field should fertilizer trees be planted?

1 = ON THE RIDGES BETWEEN PLANTING STATIONS OF THE MAIN CROP
2 = PLANTED TOGETHER WITH THE MAIN CROP ON SAME PLANT STATION
**3 = IN THE FALLOWS( BETWEEN RIDGES) AT 90CM PLANT SPACING**

D19 How many weeks after planting the main crop should you plant fertilizer trees?

**Answer: 1 week**

**Health and safety**

D20 In which direction should you face when spraying chemicals:

**1 = AWAY FROM THE WIND**
2 = TOWARDS THE WIND
3 = ALWAYS NORTH