Baseline report
ISFM Malawi

Annemie Maertens and Hope Michelson

with

Vesall Nourani and Antonia Schwarz

30 April 2015

Contents

1. Introduction

Agriculture accounts for 35% of Malawi's GDP and employs 90% of the rural population. However, 50.7% of the population engaged in agriculture lives below the poverty line. Low agricultural productivity has been attributed to: dependence on rainfed farming, low uptake of improved farm inputs, high transport costs, inadequate farmer organizations, insufficient extension services and incomplete credit, input and output markets. In response, the Malawi Growth and Development Strategies for 2011 – 2016 identified agriculture and food security as one of nine priority sectors aiming at increasing agricultural diversification, sustainable land and water management, access to inputs, extension services and markets and strengthening farmer institutions. In 2006, the Govt. of Malawi implemented a targeted farm input subsidy program. While this program has increased food production, primarily maize, the program has been criticized for its limited links to extension services and output markets.

In contrast, the Anchor Farm Model (AFM) of Clinton Development Initiatives (CDI) uses a multipronged approach. Established in 2008, the AFM is designed to increase agricultural production, income and food security through promotion of the adoption of yield-enhancing integrated soil fertility management practices (ISFM) - and soybean production in particular - by smallholder farmers in central Malawi. To reach this goal: (i) CDI disseminates production knowledge through the use of demonstration plots, farmer clubs, lead farmers and farmer field days; (ii) CDI improves farmers' access to input markets, in particular credit and seed markets through CDI's contract with seed companies and intermediary role in the credit market; (iii) CDI provides access to structured output markets through its established relationship with international soybean buyers and offers credit and storage at the time of harvest through a warehouse receipt system. AFM aims to reach 100,000 smallholders in Malawi and to expand to Tanzania by 2016 (see http://www.clintonfoundation.org/our-work/clinton-development-initiative/programs/anchor-farm-project).

In this research study, we support CDI in this scale-up. We use the random assignment of 250 villages into (i) control, (ii) warehouse receipt system and (iii) demonstration plots, to establish impacts of these interventions on farmers' income. Using detailed panel data, we identify the channels through which impacts take place and explore heterogeneity across household composition, underlying time preferences and soil properties. The latter allows us to draw lessons for SSA with the goal of increase ISFM adoption across the continent.

The research study runs from 2014 until 2019. In this report, we give an overview of the data collected as baseline in the Fall of 2014. We cover descriptive statistics, test for difference at baseline between treatment and control villages and present revised power computations. Note that this report only covers quantitative social science data collected as baseline. For a report on the qualitative data collected at baseline, refer to the qualitative and focus group report. Details on the agronomic data collected are included in the agronomic report.

2. Sample and randomization

In this section, we give an overview of the sample and randomization strategy. The data presented in this section were collected, together with CDI, while the program was ongoing, i.e., this is program participation data and not based on the data we collected during our September-October 2015 survey.

Table 1 presents the number of villages sampled in the two EPAs (Extension Planning Areas). We covered Chibvala EPA in Dowa district and Mtumthama EPA in Kasungu district. The total number of villages in these two EPAs amounts to 360, with just 303 villages with more than 50 households. We randomly selected 250 villages from these 303 villages of adequate size, and divide this set – again randomly – into two groups: treatment and control.

Within each village, we sampled 10 households from the household census as constructed by the survey team. As further detailed below, in the treatment villages where clubs were present, this sample was further stratified into 5 CDI club members and 5 CDI non-club members. If the treatment village had multiple CDI, one club was chosen at random and 5 members of that club included in the household survey.

Table 1: Number of villages in the two EPAs in treatment/control

EPA	In CDI Treatment	Control Villages
Chibvala	67	67
Mtunthama	58	58
Total	125	125

All residents of the treatment group villages were invited to a sensitization meeting run by CDI – our implementation partner – in September. In this sensitization meeting, residents of these villages were introduced to the CDI program, and were asked to form farmer clubs – at a village level with anywhere between 10 and 20 members. Table 2 gives an overview of the percentage of villages who had at least one representative attending this sensitization meeting. A detailed transcript of the sensitization meeting is available from the research team on request. We can see that the large majority of the treatment villages had someone present at the sensitization meeting while relatively few of the control villages attended.

In most cases, the individual present would be known locally as a "lead farmer or section lead farmer", i.e. a farmer who typically owns more land than the average farmer in the village, and is also known for his/her tendency to be keen on trying out new agricultural technologies. These lead farmers are often primary points of contacts for government extension agents and local NGOs. A section lead farmer is a lead farmer who is (informally) responsible for several nearby villages, i.e., for a section (the administrative unit in Malawi under the EPA level).

Table 2: Percentage of villages in the sample that had representation at the sensitization meetings by treatment/control group

EPA	% of Villages Attending Sensitization from Treatment Group (N)	% of Villages Attending Sensitization from Control Group
Chibvala	%74.6 (50)	%17.9 (12)
Mtunthama	%60.3 (35)	%15.5 (9)
Total	%68.0 (85)	%16.8 (21)

Following the sensitization meeting, clubs were formed in the treatment village. Table 3 reports the results of this process. Note that the number of villages with clubs is larger than the number of villages who attended the meetings. This is due to treatment-to-treatment spillover effects in terms of information flows, as well as an explicit effort of the CDI officer to reach the treatment villages who did not attend the sensitization meetings. Each club was then to select a leader, often this was a lead farmer in the village.

Table 3: Percentage of villages that have clubs formed by treatment/control group

EPA	% of Villages with clubs - Treatment (N)	% of Villages with clubs - Control (N)
Chibvala	%65.7 (44)	%7.5 (5)
Mtunthama	%74.1 (41)	%20.7(12)
Total	%69.6 (85)	%13.6 (17)

Several of the village who had formed clubs had in fact formed more than one club. Specifically: 27 villages had formed 2 clubs, 17 villages 3 clubs, 8 villages 4 clubs, and one village even formed has 5 clubs.

We note some treatment-to control villages spillovers effects in Tables 2 and 3. In addition, there was some spillover at the club-level where some treatment clubs included members from nearby control villages. Table 4 documents this phenomenon in Column (2): In Mtunthama, There were 5% of treatment villages that had clubs which also included members from control villages. Note also that a significant percentage of the Mtumthama clubs have out of village club members: these could include individuals from other treatment villages in our sample, control villages in our sample, or from villages outside of our sample.

Table 4: Treatment villages with club members from outside the village

EPA	% of treatment villages with club members from outside (N)	% of treatment villages with club members from control villages (N)
Chibvala	%0 (0)	%0.0 (0)
Mtunthama	% 34.9 (15)	%4.7 (2)
Total	% 17.2 (15)	%2.3 (2)

Note: This is as a percentage of the treatment villages that have clubs

While CDI was unable to change club composition at that stage of the project, CDI was able to address the first level of spillovers: from treatment villages to control villages at the next stage of the treatment: when the clubs in the treatment group were invited to their first training session. This first training session took place in December 2015. In this training, CDI gave a general introduction to various ISFM technologies and also provide basic market related information. Again, a transcript of this meeting is available on request from the research team. Table 5 gives an overview of attendance at this meeting.

We spot that attendance was significant from the treatment villages. However, we also spot that almost several of the control village clubs had some sort of presence at the first CDI training. This is due to the fact that in Mtumthama, often section lead farmers were invited as well, regardless of whether these section lead farmers came from treatment or control villages.

Table 5: Attendance to the first CDI training

EPA	% of villages with training attendee - Treatment (N)	% of villages with training attendee - Control (N)
Chibvala	%55.2 (37)	%6.0 (4)
Mtunthama	%58.6 (34)	%19.0 (11)
Total	%56.8 (71)	%12.0 (15)

Table 5: Attendance to the first CDI training (cont.)

EPA	Total number of training attendees - Treatment	Total number of training attendees - Control
Chibvala	54	6
Mtunthama	93	18
Total	147	24

Around the same time, in December 2014, CDI selected 20 demonstration plot sides in 18 of the sample villages. Later on, one side was deemed unsuitable, leaving us with 19 demonstration plots in 17 villages. Table 6 gives an overview of demonstration plot location. We note that one of the villages selected for a demonstration plot site was, in fact, a control village in our sample. The other 16 villages selected were effectively treatment villages.

Table 6: Number of villages with demonstration plots

EPA	Number of Villages with Demonstration Plots	Number of Demonstration Plots
Chibvala	8	8
Mtunthama	9	11

Once the demonstration plot sites were selected, one of the farmer clubs in the village was invited to become in charge of the plot, meaning, they would be in charge of day-to-day maintenance of the plot. Details on the demonstration plot layout are included in the progress report.

The demonstration plots were planted in January 2015. The CDI officer provided on-site guidelines for planting. Table 7 gives an overview as who attended these activities.

Table 7: Attendance to the planting activity

EPA	Number of clubs attending planting meeting – Treatment (N club members)	Number of clubs attending planting meeting – Control (N club members)	Number of other individuals attending planting meeting
Chibvala	8 (103)	0 (0)	15
Mtunthama	9 (73)	2 (25)	0
Total	17 (176)	2 (25)	15

We note that even though it was the intention of CDI to invite all nearby clubs in treatment villages to this meeting, few actually made it (as reported in column (3) of Table 7). As noted in the progress report, this was due to a multitude of factors such as delayed rainfall, short notice and the lack of a credit provided in the 2014-15 season. This issue was later on addressed by CDI when inviting club members – outside of demonstration plot villages – to other demonstration plot site activities such as farmer field days. The research team is currently in the process of collecting program participation data on the latter.

Finally, note that 2 clubs in Mtumthama from the control group were present, this is due to the fact that one control village was selected as a demonstration plot site, and this village had a total of 2 demonstration plots.

3. Descriptive statistics

In this section we present the main descriptive statistics as collected in the qualitative baseline survey in September-October 2014. We start by discussing the characteristics of the villages, and then move on to household demographics, wealth status and finally include detailed agricultural statistics including yields, adoption of ISFM technologies prices, income and perceptions.

3.1 Village Characteristics

We visited a total of 250 villages in Mtumthama EPA (In Kasungu district) and Chibvala EPA (in Dowa district) in Malawi. Among these 125 were randomly selected as treatment villages and 125 as control villages. Assisted by an enumerator, we asked the village head or village group head to complete a village questionnaire for their village

Table 8 provides some basic summary statistics at the village level. The surveyed villages consist of, on average, 76 households and 404 individuals. On average, 40% of individuals are male.

Figures 1 and 2, respectively, plot the distribution of the number individuals and households in the village. While the bulk of villages have below 1000 members, six percent of the sample villages have more than 1,000 and one percent more than 2,000 residents.

The surveyed villages are remotely located. Tables 8 to 11 give an overview. On average, villages are 3 km from a paved road and 8km from a national highway. Village residents have to travel on average between 3 and 16 km to the nearest market to sell or purchase agricultural outputs and inputs. Due to the lack of fast transportation, travel times on average lie between 36 minutes to 1 ½ hours one way. The relative remote location of village is also visible in the number of visits by government representatives or NGOs in the past year (Table 12). While NGOs pass by the villages, on average, 12 times a year, official government representatives on average only visit six times a year.

The most frequent type of organization, in which village members participate are organizations involved in agricultural financing (Table 13). Forty percent of villages have minimum one organization for this purpose. Thirty percent of the surveyed villages also have an organization for the purpose of information dissemination. There were 50 villages without any active organization.

Soy cultivation is not uncommon among the surveyed villages. However, the area used for soy cultivation and the percentage of farmers cultivating soy differs substantially among villages. Figures 3 and 4 plot the percentage distribution of cultivated area and farmer participation in soy cultivation, respectively. Table 14 provides the corresponding summary statistics. For the year 2014, the respondents estimate that about 42% of village farmers were cultivating soy and about 24% of the cultivated area in villages was used for soy cultivation.

Table 8: Composition of Villages

	Mean	Standard deviation
# of Households	75.5	77.1
# of Individuals	404.8	393.2
Avg. Male population share	40.3	8.1
Avg. Female population share	59.7	8.1

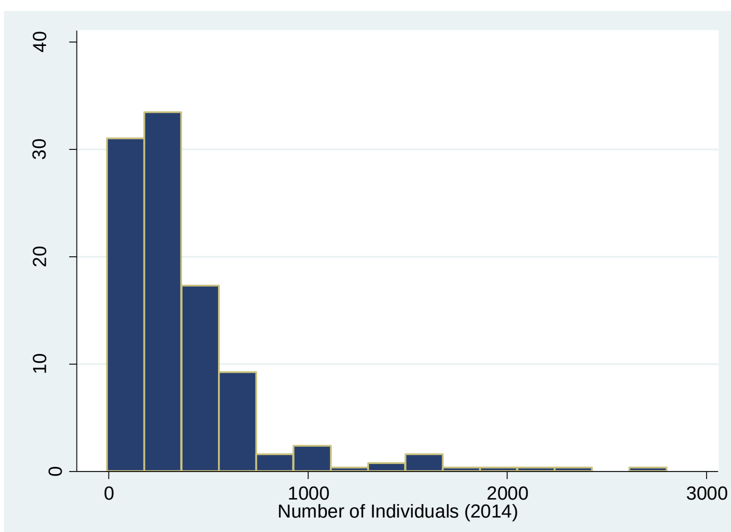

Figure 1: Number of Individuals per village

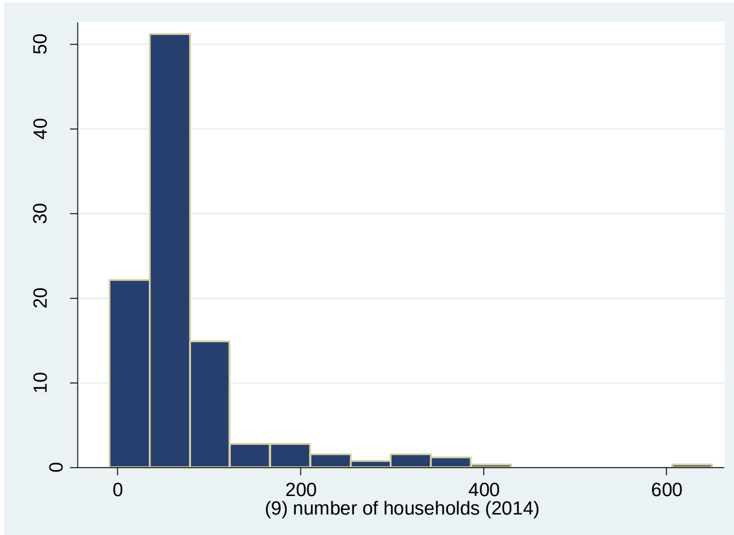

Figure 2: Number of households per village

Table 9: Distance to road in Km

	Mean	Standard deviation
Distance to paved Road	3	4.9
Distance to national highway	8	10.0

Table 10: Distance to selected location in Km

	Mean	Standard deviation
Market for pesticides, fertilizer and seeds all year	13	10.4
Market for pesticides, fertilizer and seeds during season	6	6.2
Bank or formal credit organization	16	11.0
Distance to closest market to sell produce	5	8.0
Market for fresh fruits and vegetables	3	4.1

Table 11: Travel time to selected locations in minutes

	Mean	Standard deviation
Market for pesticides, fertilizer and seeds all year	86	55.6
Market for pesticides, fertilizer and seeds all year	60	50.8
Bank or formal credit organization	90	55.1
Distance to closest market to sell produce	51	48.5
Market for fresh fruits and vegetables	35	37.0

Table 12: Number of visits by government or NGO extension agents in the past year

	Mean	Standard deviation
# of Visits by Government	6	14.7
# of Visits by NGO's	17	27.9

Table 13: Active organizations

	Avg. size (number of members)	Number of villages (out of 250)
Seed multiplication	21.7	5
Marketing of output	26.3	7
Information sharing	19.2	74
Agricultural credit	11.4	100
Labor sharing	20.0	1

Table 14: Soy cultivation

	Mean	St. Dev.
% Area under soy bean cultivation 2014	24	18.2
% Farmer cultivating soy bean 2014	42	26.8

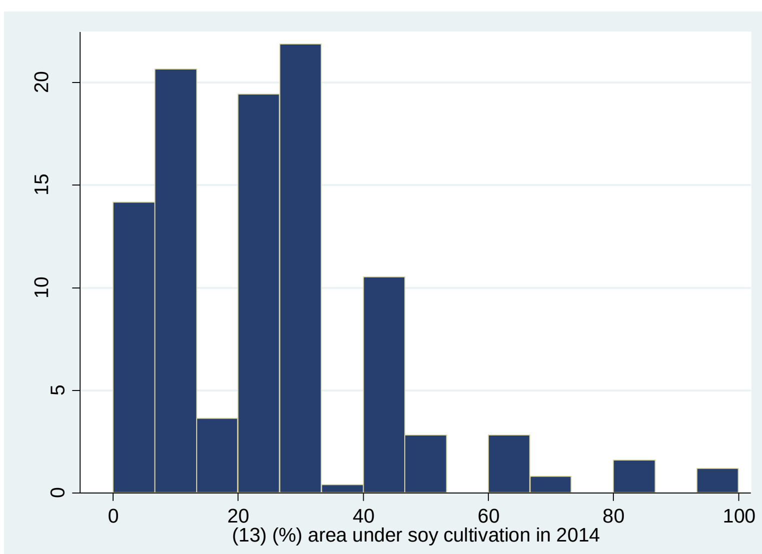

Figure 3: Area under soy bean cultivation in 2014 (village level data)

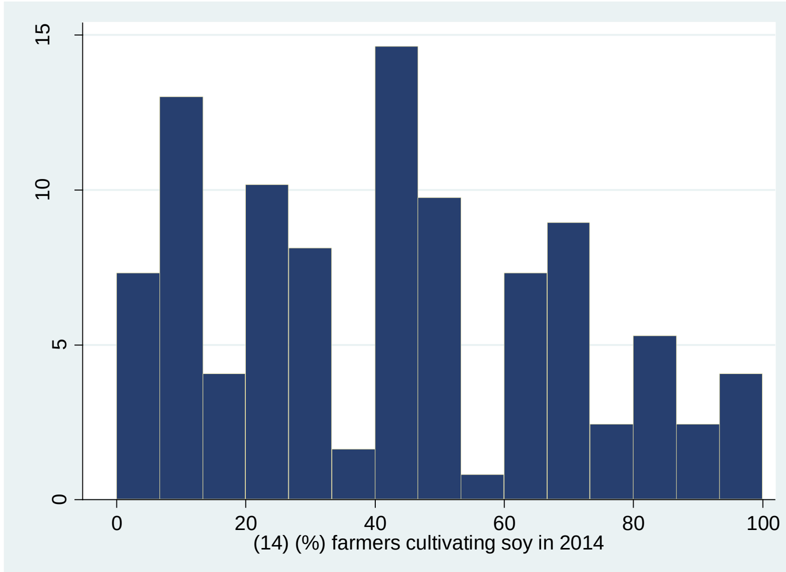

Figure 4: Farmers cultivating soy beans in 2014 (village level data)

3.2. Household demography

We defined a household as a person or group of people living together as a single unit, in the sense that they have common housekeeping arrangements. These individuals can be related to one another but need not be. More specifically, we require household members to eat at least one meal together each day. The household dataset consists of 2,501 sampled households: 1,250 households in treatment villages and 1,251 households in control villages.

In each household, we interviewed one member. In 70% of the cases, we interviewed the household head. In 30% of the cases, we interviewed another knowledgeable member of the household.

Table 15 presents some information on the household head. Household heads are on average 42.4 years old and have received, on average, 6.1 years of education. About 18% of household heads are female.

Table 15: Characteristics of the household head

	Mean	St. Deviation
Gender household head (1=female, 0=male)	0.18	0.39
Education level (years)	6.1	3.48
Age (years)	42.4	15.19

.

Through a household composition module, we collected basic information about a total of 12,850 individuals. Table 16 to 18 presents some selected descriptive statistics.

The average size of households in the sample is 5 members with, on average 2.7, children per household (defined as being between 0 and 18 years). More than 80% of the individuals in the dataset are either respondents or directly related to the respondent (i.e., wife/husband or children; see Table I in the Data Appendix). About 50% of individuals are male. More than 60% of individuals age 18 and above are married in a monogamous relationship.

Table 18 reports the highest level of education obtained for members over 18 years. About 84% of individuals have obtained some form of education. More than 60% of individuals aged 18 and above have obtained 1 to 5 year of education. More than 30% finished between 6 and 8 standards. In Malawi, 8 standards is equivalent with primary education. The next four years are known as secondary education. A further 10% have an additional 2 years and another 10% obtained an additional 4 years. Tertiary education is limited with less than 1% of the sample having some form of tertiary education.

Table 16: Household composition

	Mean	St. Dev.
# of members	5.1	2.2
# of children	2.7	1.8

Note: Children are defined as age 0 to 17.

Table 17: Marital status (for age 18 and above)

Marital status	Percent of individuals
Never been married	20
Married/monogamous	62
Married/polygamous	7
Separated	3
Divorced	4
Widowed	5

Table 18: Education level (for age 18 and above)

Highest level achieved	Percent of individuals
None	15.4
Pre-school/nursery	0.2
Adult literacy	0.7
Standard 1-5	31.7
Standard 6-8	31.8
Form 1-2	9.9
Form 3-4	9.9
Some tertiary	0.4

Table 19 reports the primary activity of individuals age 18 and older. This is the main activity the individual was engaged in over the last 12 months (in terms of time contribution). Most individuals are engaged in agricultural activities on their own farm (over 70%). A further 10% of adults primarily go to school. About 8% of adult individuals are involved in some business activities. Less than 3% reported wage employment as their primary activity.

Table 19: Primary activity (for age 18 and above)

	Percent of individuals
Education	10.9
Household activities	5.7
Agricultural activities on own farm	70.6
Business activities	7.7
Wage employment-private company/individual	2.0
Wage employment-govt/estate owned enterprise	0.9
Masaf/public works program	0.1
Ganyu (daily labor) on farm	1.0
Ganyu (daily labor) non-farm	0.5
None	0.7

3.3 Social networks and CDI participation

The prevalence of the agricultural sector in the daily lives of the households is also reflected in organizational participation. In total, 45% of households participate in at least one organization. Conditional on participating in at least one organization, on average, each household is active in 1.3 organizations.

Excluding CDI club membership, Table 20 shows that over 35% of households are a member of a farmer's group and about 28% of households have at least one active member in a trade union.

When asked about the usefulness of these organizations, in 60% of the organizations mentioned by the households, the respondent reported the organization to be very useful for them while for 33% of the organizations mentioned, the respondent thought they were not working to their benefit.

The latter could be related to the infrequent meetings of these organizations (see Table 21). Fifteen percent of households report the organization to have met not once in the past 12 month. However, for over 50% of the organizations mentioned, the respondent reported they met more than one per week.

Table 20: Participation in organizations

	Percentage of households
Farmer's group	35.02
Trade union	27.66
Youth group	3.28
Parent group	1.24
School committee	2.66
Sports group	1.15

Note: organizations with over 1% participation rates were included in this table

Table 21: Frequency of meetings

	Percentage of organizations listed that meet…
More than once/ week	15
At least once/week	38
At least twice/month	14
At least once/month	9
Less than once/month	5
Never	15

Recall that members of the treatment village were invited at the CDI sensitization meetings. When asked about familiarity with the CDI program, about 50% of the respondents in treatment villages stated that they had heard about the program (as opposed to just 10% in the control villages). As the survey happened very soon – days to weeks – after the sensitization meeting took place, one could expect the percentage in the treatment village to have gone further up by the time we'll collect the midline data.

When asked about from whom they have heard about the program, about 36% points at fellow farmers, while close to 25% heard about CDI directly from either the CDI agent or the government extension agent.

Twelve percent of households report to be a member of a CDI club. This is naturally markedly different between treatment and control, only one household in a control village reports to be a member of a CDI club. Note that this 12% is markedly lower than the expected 25% (in the treatment villages, half of the households sampled is a club member), and reflects the fact that not all treatment villages have formed clubs. In effect, the baseline data reveals that there are 37 villages without any (self-reported) CDI villages. Sometimes, several members of a household can be a CDI club member. There are 16 such households in the sample.

In terms of contacts outside of the village, 78% of respondents report that they make more than once per week trips to one of the neighboring villages, and 65% reports to make more than once per week trips to a nearby trading center. However, only 40% of the respondents go at least once per month to the district capital.

Television does not appear to be a major source of information for the respondents: more than 85% states that they never receive information about agriculture from TV. Neither is the internet important, with close to 100% of the respondents stating they never receive information through the internet. The radio is a different matter, and over 50% states that they often receive information about agriculture through the radio. In terms of personal contacts, only 10% of the respondents claim to receive information from government extension agents on a regular basis, while over 40% states they never receive any information from them. Over 35% of respondents sometimes receive information from agro-dealers and other private agents – but over 60 % states they never receive information from these sources. NGOs are about equally important, with over 60% of respondents stating that they never receive information from this source.

When asked the open question about whom is most important to them in terms of agricultural information (up to 3 answers were allowed), most respondents point to fellow farmers in the village (about 60% of contacts mentioned were fellow farmers). Lead farmers and government extension agents share the second position, with about 5% of individual mentioned lead farmers, and about 20% government extension agents.

In addition, households in treatment villages were asked about their familiarity with members of the sampled farmers' club in the village (recall that one club per treatment village was sampled, i.e., 5 members of that club were included in the sample). Table 22 provides an overview. We read out a list of all (often 20) club members to each respondent (in the treatment village). Let us denote the name of each club member by X. We then proceeded to ask whether the respondent knew X and if so what the nature of their relationship is. Table 22 reports the results.

Eighty-two percent of club members were known to respondents. Among these, the respondent had asked 21% for advice on farming in the past. The respondent, respondents feel confident to approach more than 80% of the (known) club members and would entrust a valuable item to about 70% of the (known) club members.

Table 22: Familiarity with CDI club members in the treatment villages

[As a %age of matches X]

	%
Do you know this person?	82
If yes, have you asked this person for advice on farm activities?	21
If yes, could you approach this person for advice?	88
If yes, would you trust this person to look after a valuable item?	73

3.4 Time preferences

We obtained time preferences through a hypothetical exercise in which we asked the respondent their Willingness-to-Pay (WTP) for an agricultural technology which would yield benefits only three years from now, but for which they would have to pay right now. The technology imagined would increase maize yield to 80 bags/acre. The hypothetical price ranged from 100,000 Kwacha (about 250 USD at the time of the survey) all the way down to 0 Kwacha. Once the respondent answered they were willing to pay an amount X, the exercise was stopped and this value X was taken as their individual WTP. The protocol of this exercise is available on request from the researchers.

Figure 5 reports the results.

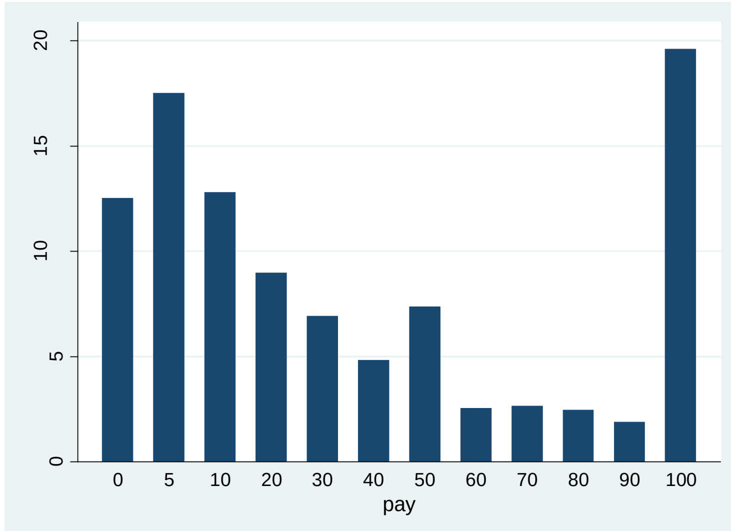

Figure 5: Willingness to Pay

A large number of respondents (over 20%) were willing to pay the maximum amount, while over 10 % of respondents were not willing to pay anything at all. The non-normal shape (but in fact bi-modal) of the distribution in Figure 5 makes us suspect that we at bumped into credit-constraints for the poorer households, while for the richer households, the fact that the investment would yield a per acre – and not a total – yield would have made our maximum price – 100,000 KW – sound like peanuts indeed.

We further explored this hypothesis by investigating the histogram by acreage of land owned. We see that among poorer households, a WTP of 0 is indeed more common, while for the richer households, a WTP of 100,000 is more common (These results available on request). We hence propose to repeat this exercise in the midline survey, this time presenting the exercise as a more scale-neutral investment.

3.5 Landholding and assets

3.5.1 Landholding

We asked the household head to list all the fields they owned in the 2014-15 season in addition to all the cultivated fields in the 2013-14 season. By season we referred to two seasons: the rainy season and the dry season. Malawi has one rainy season, running from December till May. The fields cultivated in the rainy season are referred to as Munda fields, while the fields cultivated – typically through irrigation next to the river - over the dry summer spell are called Dimba fields.

Over 97% of households reported owning some land. In fact, landownership is rather common, and over 88% of the fields mentioned by the respondent are owned (as opposed to rented or leased in). There are only few fields reported to be communal land.

The average number of fields listed is 2 per household (standard deviation is 1.03), with a maximum of 7 fields listed. When asked about who made decisions regarding this field, for over 80% of the fields a male household member was listed.

The most common type of fields are Munda fields (> 80% of fields in our sample). Just 16% of fields in our sample are Dimba fields. The size of the Munda fields is, on average, 2.4 acres (with a standard deviation of 2.3). The size of the Dimba fields is, on average, 1 acres (standard deviation of 1.25). (Note we dropped outliers over 1% for this computation)

We asked the respondent to 'guess' the sales value of his/her owned plot. While this question does not appear to always have been properly understood, it is interesting to note that the Dimba fields, while smaller, appear to be considered more valuable. Their sales price is estimated at, on average, 210,250 Kw (standard deviation of 250,608 Kw) while the sales price of the Munda fields estimated at, is, on average 173,430 Kw (standard deviation of 226,822 Kw). Again, we dropped outliers above the 1% quartile for the latter computations. Most of the owned fields are acquired through inheritance. A few fields are purchased or granted by a local leader. A minority of fields come as a bride price or are illegally obtained.

In terms of soil texture, 43.7% of fields have a mixed soil, 35% are clay, 17.3% are sand, and 0.4% an alternative soil composure. Table 23 provides an overview of the soil fertility, as perceived by the respondent. Over 20% of fields are perceived to be of very poor to somewhat poor soil fertility. Soil fertility seems to have worsened over the years, as reported in the answer to the question: 'In the last five years, has the soil fertility of this field…[options read out]': 80% of fields are perceived to have either stayed the same or worsened in terms of soil fertility. Independent of the soil type, most farmers experienced soil degradation in 2014, especially in form of nutrient degradation but also soil erosion. Specifically, 43% of fields are reported to suffer from soil erosion, 53% suffers from nutrient depletion, and 22% suffers from water logging. A simple cross tabulation reveals that the quality of fields greatly varies by soil type. Clay fields on average have the best quality, with 60% being above average (Results are available on request)

Table 23: Perceived soil fertility status

Question	Question options	Percentage of fields
What is the soil fertility of this field? (yes)	Very poor	6
	Somewhat poor	17
	Average	25
	Somewhat Good	31
	Very Good	21
In the last five years, has the soil fertility of this field (yes)	Improved a lot	6
	improved a little	12
	stayed the same	40
	became worse a little	37
	became worse a lot	4
Does this field suffer from soil degradation in the form of (multiple)	Soil erosion	43
	Nutrient depletion	53
	Water logging	22
	Salinity/acidity	5

3.5.2. Assets

We computed an asset index using a principal components analysis (PCA). The first principal component is the n-dimension eigenvector that captures the most information common among the assets, the linear combination of the n-variables with maximum variance. The use of the first principal component as a vector of asset weights assumes that household long-run wealth is the common factor underlying the component of maximum variance in the data. The PCA asset index serves as a relative measure of poverty for each household, meaning that the household's poverty ranking is made relative to all the households used in the PCA weights estimation.

Household characteristics and assets included in the PCA include household durable assets as well as household characteristics including roof and floor material, source of lighting, water, and sanitation. In addition to this index, we constructed a second measure of assets based on reported asset values. The distribution of this aggregate reported asset wealth is presented in Figure 6.

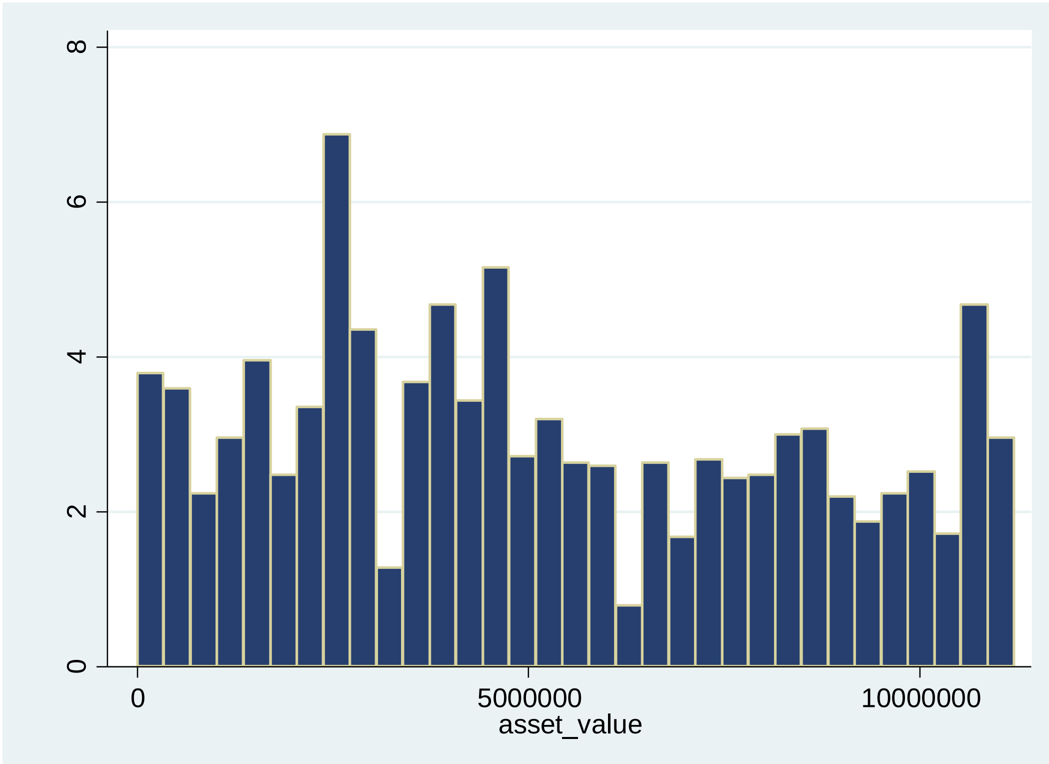

Figure 6: Distribution of total asset wealth [in Kw]

3.6 Participation in input and output markets

Eight-five percent of households report participating in the agricultural output market in the 2013-14 season, i.e., these households have sold or plan to sell some of their output of the 2013-14 season.

Market participation greatly varies by crop. Table 24 gives an overview for the most important crops in the area: Soya beans, Maize, Groundnut, Beans and Tobacco.

Table 24: Market participation – by crop

	Market participation % of households who sold or plan to sell	Total amount sold/planning to sell as a percentage of total harvest
Tobacco	99.1	99
Soya bean	87.3	75
Groundnut	66.4	43.5
Beans	54.2	37.1
Maize	44	15.6

Note: This percentage is for all households who have harvested this particular crop

Farmers sold their crops mostly in the village, either to a farmer in the village or to a trader/commission agent who visited the village. Table 25 gives an overview – again for the most important crops. One can see that the marketing channels are very much crop dependent, with almost 60% of the soya beans sold outside of the village versus less than 10% of the maize.

Table 25: Marketing channels – by crop

	Soya beans	Maize	Ground nuts	Beans	Tobacco
To a farmer in the village	4.6	14.5	23.3	16.4	17.7
To a trader/commission agent who come to/lives in village	33.3	75.3	68.0	76.3	77.6
To an NGO who came to village	0.6	0.7	0.4	1.2	0.0
In an agricultural market	2.3	1.0	1.0	0.1	0.7
To a trader/commission agent in nearby town	57.0	8.4	7.0	5.8	4.0
To an NGO in nearby town	2.1	0.0	0.0	0.1	0.0

Farmers also participate in the credit market. 20% of farmers report receiving some kind of loan for the 2013-14 season. When asked about constraints to credit access, farmers report that they could approach friend or relatives (35%), village credit groups or associations (14%), or a micro-credit lender (30%). Few farmers – about 10% mention a formal source of credit, such as a bank. Almost 50% of farmers believe that if they were to approach this abovementioned source for credit, they would be likely to receive it.

Farmers in Malawi also receive assistance from the government through the well-known Farm Input Subsidy Program, established in 2005-06 to improve small farmer maize productivity. In effect, 54% of the households in our sample report having received some kind of subsidy assistance from the government.

3.7 Beliefs and adoption of ISFM techniques

We have three sources of information regarding beliefs and adoption of ISFM techniques. First, we elicited the perceived yields on an per acre and per field basis for soybean, groundnut and beans. We used the following script: 'Imagine that you would cultivate maize this coming year, and imagine that maize would be the only crop on the field, so no other crops were present, in such a case, how much maize do you think you would harvest in bags of 50 kg?'. Note that we phrased the question as such as to refer to a situation of mono-cropping as most fields (as noted in a later section of this report) are set up as mixed cropping arrangements. To check for consistency we then asked how much maize the farmer would harvest on one of their fields (where we also recorded the field acreage).

Second, we asked the respondent a series of recall questions (going back 5 years) on used ISFM technologies. Due to the recall period involved, we limited these questions to discrete yes/no type questions only. Table 26 reports the results. The majority of farmers report preparing the land by hand/hoe and using ridges to prepare the land for cultivation. A significant proportion of farmers complements the ridges with other features such as grass strips, drainage channels and water catchments with the goal of improving soil structure. Framers report using a wide variety of techniques to improve soil fertility, ranging from incorporating crop residue at the time of planting to crop rotation. Most farmers report having planted soybean in the past (over 80%). The numbers are equally large for bean and groundnut, with over 90% reporting some cultivating experience on these two crops. Few farmers have experience with the so-called fertilizer trees and farmers do not actively use forms of pesticide and insecticide.

Table 26: Use of ISFM technologies in the past 5 years

Past Farming Activities (in %)

		% of households
Have you used any of the following methods of land preparation in the past 5 years [% reports yes]	No tillage	4.5
	By Hands/hoe	99.6
	Ploughing with Animals	1.0
	Mechanized	0.2
	Other	0.2
Have you used any of the following soil conservation methods in the last 5 years [% reports yes]	Grass Strips	27.0
	Ridges	99.1
	Bench Terraces	1.2
	Drainage Channels	25.3
	Water Catchment	12.7
	Other	9.1
Have you used any of the following methods of soil Fertility Improvement in the past 5 years [% reports yes]	Crop Residue	66.7
	Animal Manure	63.4
	Inorganic Fertilizer	93.2
	Improved Fallow	18.0
	Legume Cover crop	3.5
	Compost	51.2
	Intercropping	34.8
	Crop Rotation	84.6
	Other	81.5
Have you planted any of the following legumes in the past 5 years [% reports yes]	Soybean	82.6
	Pigeonpea	12.8
	Groundnut	91.8
	Common bean	89.5
	Other	9.7
Have you planted any of the following soil fertility enhancing trees in the past 5 years [% reports yes]	Tephrosia	6.6
	Gliricidia	6.8
	Sesbania	2.8
	Other	12.1
Have you used any of the following pesticides/Herbicides in the past 5 years [% reports yes]	Insecticide	20.5
	Herbicide	1.2
	Fungicide	2.8
	Fumigant	2.1
	Other	0.1

Finally, to avoid recall bias, we also asked farmers about current use of ISFM technologies in the 2014-15 season through a plot-level questionnaire. This plot-level questionnaire was administered for each plot that was cultivated during the 2013-14 rainy season (so it includes Munda plots only). Note that each field can consist of multiple plots. A field was defined as a continuous piece of land, while a plot was defined as a section – within a field – where a particular cropping arrangement prevails.

About 75% of the plots featured one sole crop, i.e. mono-cropping, while the remainder features some form of mixed cropping, either row intercropping (23%) or strip intercropping (<2%). In both types of intercropping arrangements, the plants are arranged on ridges. Take for instance the example of maize and beans. In strip intercropping, there are alternating ridges of maize and beans; while in the case of row intercropping, maize and beans would be present on the same ridge: one maize plant could for instance follow one bean plant. In some instances, plants are sown at the same station, i.e., at the same location.

Groundnut, maize, tobacco and soybean are all commonly cultivated in both mono-cropping as well as intercropping systems.

Before we proceed, let us take a look at the most important crops among our sample households. Table 27 report the percentage of households that cultivate a particular crop, as well as the acreage, conditional on cultivation.

Table 27: Cultivated crops in 2013-14

		For all households who cultivate this crop	
	% of households that cultivates this crop in 2013-14	Mean acreage	Standard deviation of acreage
Local maize	46.0	1.0	0.88
Hybrid maize	53.3	1.1	1.14
Groundnut	53.7	0.8	0.56
Bean	43.1	0.6	0.52
Soybean	42.1	0.8	0.84
Burley tobacco	42.0	1.3	2.44
Nkhwani	5.4	0.4	0.5
Sweet potato	3.9	0.5	0.44
Composite maize	2.7	0.7	0.57
Tobacco other	1.8	1.5	1.3
Cassava	1.6	1.2	1.27

Tables 28 to 30 report these plot-level results. Table 28 summarizes the results which pertain to land preparation. Again, most plots used ridges and are prepared by hand/hoe. Crop residues were incorporated in 10% of the plots and about 13% of plots received some form of animal manure.

Table 28: Land preparation techniques used in 2013-14

	Percentage of plots
Grass strips	12.0
Pits	1.0
No tillage	0.9
Application of animal manure	12.8
Ridges	99.6
Water catchments	3.2
Tillage by hand/hoe	98.7
Application of crop residue	9.1
Marker ridges	17.7
Drainage channels	9.6
Ploughing with animals	0.4
Other biomass	0.7
Box ridges	54.3
Permanent trees	14.8
Ploughing by tractor	0.2

Table 29 shows the use of hybrid/improved seeds and seed treatment. It appears that farmers might be unfamiliar with the concept of seed inoculation – a technique used in soy farmer to enhance the performance of soy seeds. Farmers appear to be very familiar though with improved seeds and use them on 53% of their plots.

Table 30 summarizes the use of inputs which improve soil fertility. While non-organic fertilizer is commonly used, with almost 97% of the plots using some, the use of organic fertilizer – at least during the growth cycle is less common. These fertilizers include animal and human manure and compost, among others. Informal conversations with respondents and extension agents revealed that transportation issues might be a key constraint here: without a wheelbarrow of some kind, it is difficult to get these organic fertilizers to sometimes far away plots. Even though 85% of all fields are within a one hour walking distance of the residence of the respondent, the fact that fields are often scattered in multiple directions can make travelling a challenge.

Table 29: Seed and seed treatment 2013-14

	Percentage of plots
Use of hybrid seed	53
Seed Treatment	0.5

Table 30: Use of fertilizer, crop residue and fertilizer trees in 2013-14

	Percentage of plots
Use of non-organic fertilizer	97
Use of organic fertilizer	9
Fertilizer trees planted	5
Crop residue incorporated or plan to incorporate in soil	48

Finally, Table 31 reports the descriptive statistics on farmers' perceptions regarding the expected yield of soy, groundnut and maize. We selected these three crops both because they are important crops in the Malawi context, but also because they are the focus crops (in addition to the common bean) of the CDI Anchor Farm Program. Figures 8 through 10 present the respective distributions.

The median yield distribution of maize is 25 50kg bags. The median groundnut yield is 15 50kg bags. The median soybean yield is 10 50kg bags.

Comparing these perceived yields with the actual yields reported for the 2013-14 season, as reported in Table 34, it seems that farmers have a good idea of what to expect in most cases, but might be a little optimistic in some. The perceived median maize yield is a little above the actual average maize yield for the sample (the actual is 10 to 15, depending on type). The median perceived groundnut yield is again a little above the actual yield for the sample (the actual is 10). The only number which stands out here is the median perceived soybean yield, which at 10 is about 5 times the actual average yield obtained on farmers' mono-cropped fields in 2013-14. We intend this issue further in future analysis.

Table 31: Perceived yield of soybean, groundnut and maize

	Median	Mean	Standard deviation
Maize [in 50 kg bags, shelled]	25	28	15
Groundnut [in 50 kg bags, unshelled]	15	20	13
Soybean [in 50 kg bags, shelled]	10	12	7

Note: we dropped outliers in the top and bottom 1 percentile.

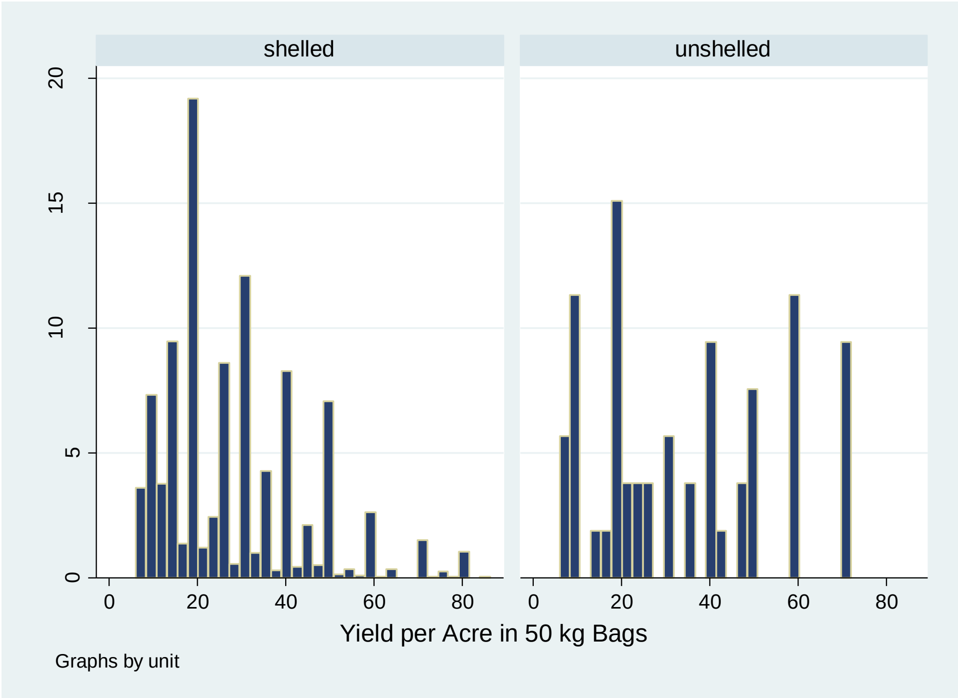

Figure 7: Perceived yield of maize [in 50 kg bags, per acre] – note only 56 farmers reported the unshelled type

Note: we dropped outliers in the top and bottom 1 percentile.

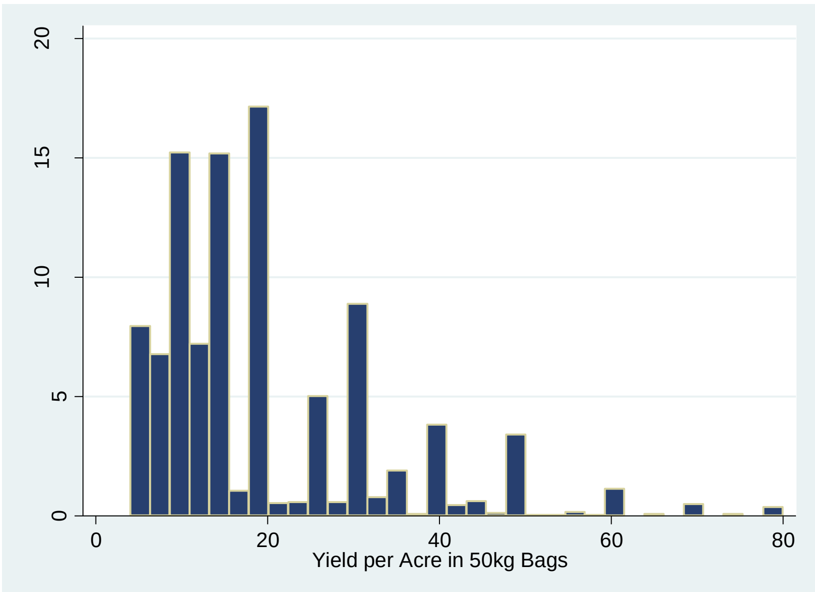

Figure 9: Perceived yield of groundnut [in 50 kg bags, per acre, unshelled, dried]

Note: we dropped outliers in the top and bottom 1 percentile.

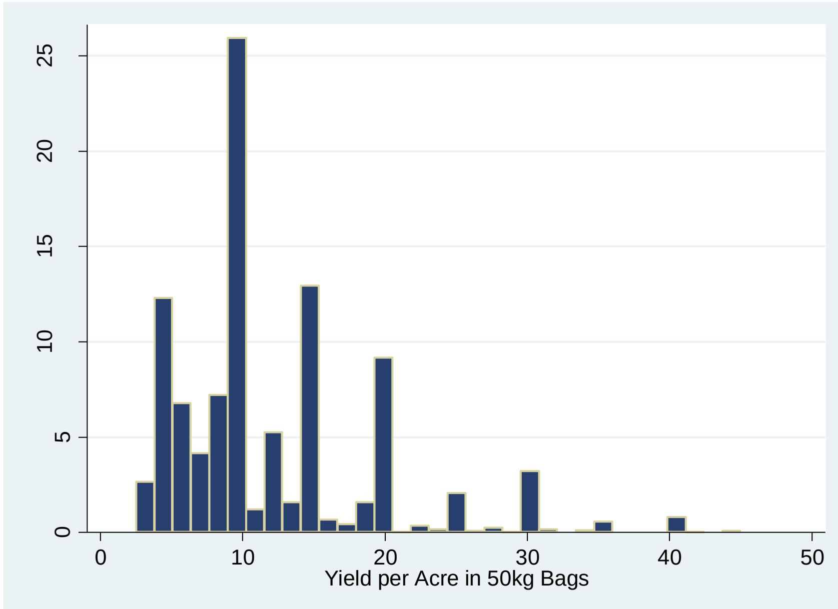

Figure 8: Perceived yield of soybean [in 50 kg bags, per acre, shelled]
Note: we dropped outliers in the top and bottom 1 percentile.

We conclude this section with a note on data quality. One issue we noted on the field is that respondents often have a difficult time to imagine the size of an acre, and hence might have difficulty in estimating per acre yields as such. To deal with this concern, we repeated the yield question, but this time asked the respondent to report their estimate for their fields specifically. Using then acreage data on their fields (which are again, arguably, subject to measurement error), we can compare the per acre estimate with the per-field estimate. Table 32 reports the results of this exercise. We note that over 70% of the respondents with their per-field estimate is within a 30% range of the per-acre estimate.

While this is not a bad outcome, in terms of precision, for the midline, we propose to (i) use the actual field area as estimated for fields that were selected for soil sampling for whom we have the factual acreage as estimated by GPS, and (ii) perhaps drop the households for whom these estimates diverge too much for the impact analysis.

Table 32: Incorrect yield reporting

	% of respondents within 30% range	% of respondents outside of 30% range
Maize	76	24
Soybean	77	23
Groundnut	75	25

3.8 Agricultural yield and income

Using the plot-level questionnaires mentioned earlier, we can compute the yield for each crop. All but 63 farmers completed at least one plot-level questionnaire. Note that this does not imply that these farmers do not have any agricultural income. Instead, this means that these farmers did not cultivate any plots themselves during the 2013-14 rainy season. We have no information on other sources of income, and hence cannot say anything about the income through rented out land, for instance.

To compute the agricultural income for the 2013-14 rainy season, we combine the yields, with information on output prices and input costs.

3.8.1 Agricultural yields

We faced two complications when computing yield per acre. The first complication is that farmers' estimate of plot acreage is usually fairly tentative, making measurement in this variable likely. While we did not have the budget to measure all of the plots using GPS tools ourselves, we did measure a subset of the fields in the agronomic survey and intend to compare these measurements with the farmers' estimates to get a sense of the size of the measurement error.

In addition, as noted earlier in this report, some plots used some form of mixed cropping system. As noted earlier, 75% of plots were cultivated under a mono-cropping system (only one crop was planted), while the rest of the plots used some form of mixed cropping system. Mixed cropping, with multiple crops per plot, creates complications for yield estimates as one needs to allocate a share of the plot to a particular plot. Asking the respondent to estimate this share directly usually does not work well, as it implies knowledge of percentages combined with a good memory of the geography for each plot. Instead, we asked the respondent to provide us with the cropping pattern. For instance, if a system of strip intercropping is applied, we ask the farmer how many ridges of crop A alternate with how many ridges of crop B. Using this pattern of ridges, we estimate the share of the plot allocated to each crop.

A last complication we faced is that the respondent can report their harvest in many different units, ranging from kg, to bales and carts. Naturally, the crop-specific conversation ratios are merely estimates, and natural variation in these conversations will introduce another level of measurement error. We provide an overview of the conversions we used in Table 33. For this reason, we report the yields in the upcoming tables always in the most commonly used unit.

Table 33: Conversion rates used for most common crops

	Unit	Condition	Conversion to unshelled	Conversion to common unit-50 kg bag for all except tobacco (pail)
Bean	50 kg Bag	Shelled	0.38	1
Bean	Kilogram	Shelled	0.38	0.017751
Maize	50 kg Bag	Shelled	0.83	1
Maize	Kilogram	Shelled	0.83	0.018765
Groundnuts	50 kg Bag	Shelled	0.69	1
Groundnuts	Kilogram	Shelled	0.69	0.02459
Irish Potatoes	50 kg Bag	NA	NA	1
Irish Potatoes	Kilogram	NA	NA	0.022472
Soybean	50 kg Bag	Shelled	0.38	1
Soybean	Kilogram	Shelled	0.38	0.016878
Sweet Potatoes	50 kg Bag	NA	NA	1
Sweet Potatoes	Kilogram	NA	NA	0.022329
Tobacco	50 kg Bag	NA	NA	0.2
Tobacco	Pail (large)	NA	NA	1
Tobacco	Pail (small)	NA	NA	0.36

As the method outlined above results in merely an estimate of yields, we first present the yield estimates for the mono-cropped plots only. Table 34 reports the results. We complement this table with Figure 11, Figure 12 and Figure 13, respectively presenting the yield distribution under mono-cropping cultivation of both hybrid and local maize, groundnuts, and soy bean; which are the main crops the CDI program focusses on.

Table 34: Per acre plot-level yields of most common crops (based on mono-cropped plots)

	Average	Standard deviation	Unit	N
Local maize	9.8	7.6	50 kg bag shelled	709
Hybrid maize	14.3	10.7	50 kg bag shelled	813
Groundnut	10.3	8.1	50 kg bag unshelled	1,217
Bean	1.9	1.5	50 kg bag shelled	60
Soy bean	2.5	2.1	50 kg bag shelled	107
Burley tobacco	3.7	2.3	bale	554

Note: We dropped outliers in the top and bottom 1 percentile. We dropped all crops with N smaller than 50.

In Table 35, we now present the equivalent results, but this time only including plots featuring a mixed cropping pattern. Comparing Tables 34 and 35, we note that the yields of both types of maize are lower in the mixed cropping systems compared to the mono-cropping systems. However, also note the relatively higher variance in Table 35, possibly due to measurement introduced by computing per-acre yields in mixed cropping systems. Hence, we are careful not to over-interpret these results.

Table 35: Per acre plot-level yields of the most common crops (mixed cropping systems only)

	Average	Standard deviation	Unit (including shelled/unshelled)	N
Local maize	5.1	5.7	50 kg bag shelled	80
Hybrid maize	5.7	6.2	50 kg bag shelled	68
Bean	0.2	0.1	50 kg bag shelled	100

Note: We dropped outliers in the top and bottom 1 percentile. We dropped all crops with N smaller than 50.

Finally, in Table 36, we combine the two sources of information and present the estimates based on all plots, mono-cropped and intercropped.

Table 36: Per acre plot-level yields of the most common crops

	Average	Standard deviation	Unit	N
Local maize	9.3	7.6	50 kg bag shelled	796
Hybrid maize	13.6	10.6	50 kg bag shelled	882
Groundnut	10.2	8.1	50 kg bag unshelled	1,236
Bean	0.8	1.1	50 kg bag shelled	161
Soy bean	1.9	2.0	50 kg bag shelled	146
Burley tobacco	3.6	2.3	50 kg bag shelled	562

Note: We dropped outliers in the top and bottom 1 percentile. We dropped all crops with N smaller than 50.

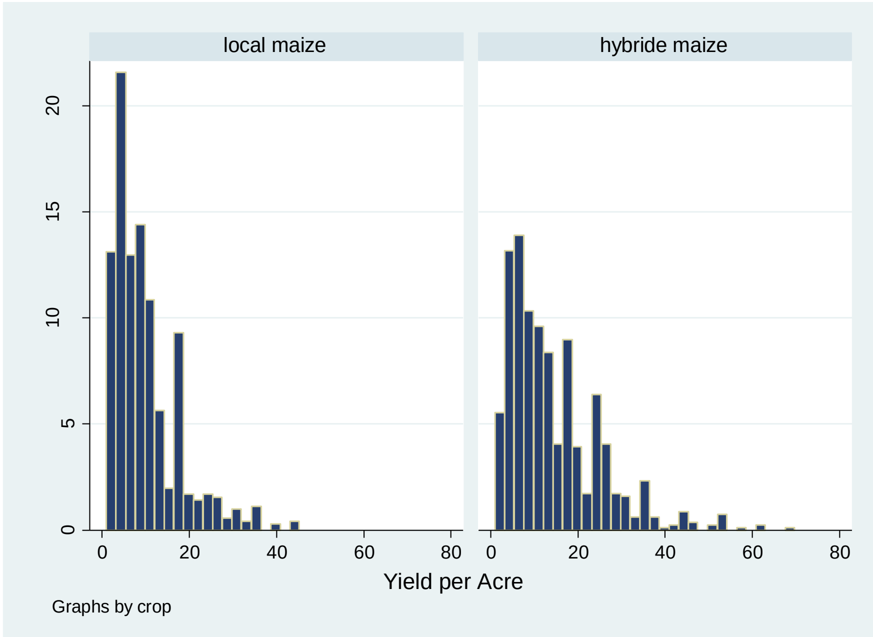

Figure 11: Yield distribution of hybrid and local maize [50kg bag shelled]
Note: We dropped outliers top and bottom 1 percentile

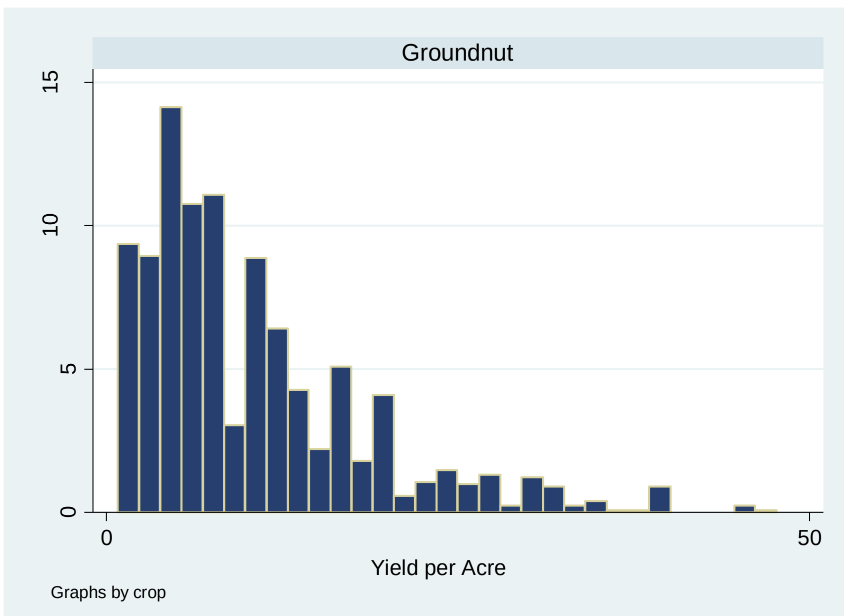

Figure 12: Yield distribution of groundnut [50kg bag unshelled]
Note: We dropped outliers top and bottom 1 percentile

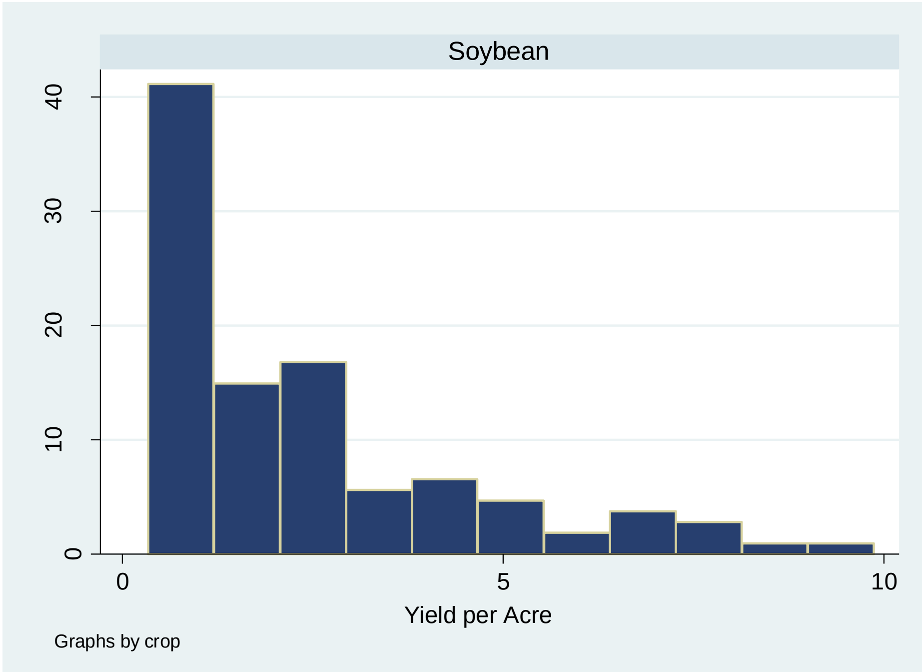

Figure 13: Yield distribution of soybean [in 50kg bags shelled]
Note: We dropped outliers top and bottom 1 percentile

3.8.2 Agricultural income

As mentioned earlier, to compute agricultural income, we combined the information on plot-level yields, input costs and output prices.

We first estimate household agricultural revenues for the 2013-14 rainy season as follows:

$$R_i = \sum_{for\ all\ k} Harvest_{i,k} * P_{i,k}$$

Where i refers to the household and k refers to the crop. Thus revenue equals the sum of revenues from each crop, where revenue of a crop is defined as the harvest amount multiplied by the village-level crop-specific price, meaning we value both sold and non-sold units. Note that not all of the harvest is sold (see also Table 24 on this matter).[1]

We then proceed and estimate the household costs for the 2013-14 rainy season by

[1] We considered using a farmer-level price. However, not all farmers sell on the market, or had sold by the time we had interviewed them. In addition, due to a discrepancy in unit reporting between prices and quantities such a farmer-level price might suffer from additional measurement error. We intend to address this issue with further data cleaning and in the midline instruments.

adding, for each plot, the various paid out costs as reported by the respondent. These costs can include costs of hired labor, pesticides, fertilizers and seeds. Note that we did not impute an opportunity cost for the costs which were not paid out, such as own family labor. As this imputation is fairly complex, and various methods can be used, we leave this computation for further analysis.

Excluding top and bottom 5% outliers, we find that the average agricultural (rainy season) household income is 108,677 Kw. This is equivalent with 271.70 US dollars at the time of the survey (400 Kw – 1 USD). The standard deviation, as expected is fairly large 102,338. The median of this distribution stands at 72,322.

Dividing this household-level agricultural income by the number of family members, we find compute the per-capita agricultural rainy season income, again dropping the top and bottom 5% of the distribution, we find that the average per capita income is 11,227 Kw and the standard deviation is 10,241. Figure 14 provides the corresponding distribution.

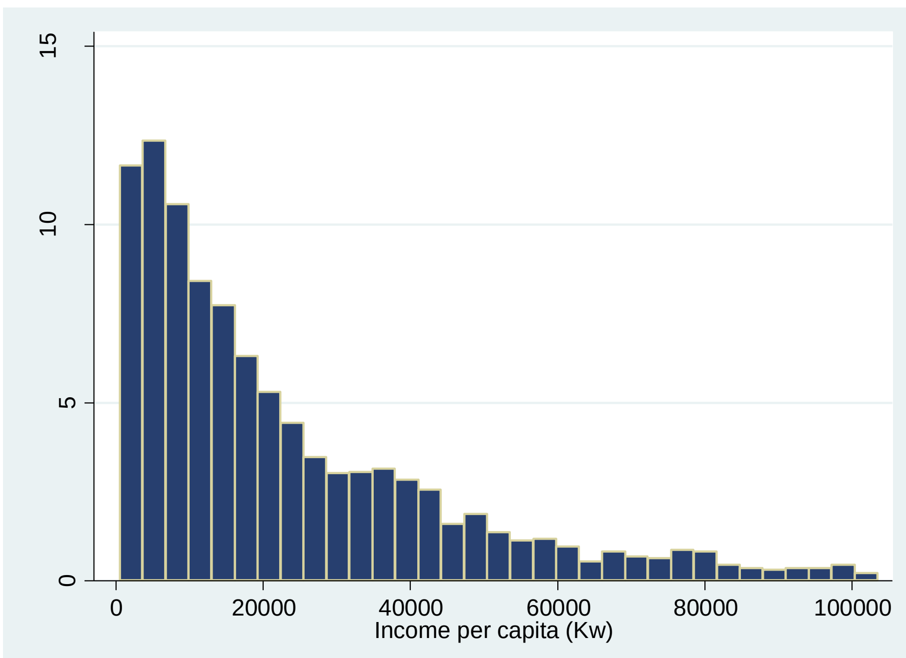

Figure 14: Distribution of rainy season 2013-14 (per capita) income [in Kw]

Note: outliers under and above 5% were dropped

3.9 Club dynamics

To obtain a basic measure of club dynamics among the newly formed CDI farmer clubs, we conducted a public goods game among one randomly selected club in each treatment village where we had a club present (Recall, not all treatment villages formed clubs).

In this public goods game, we asked each participant to split 400 Kw (at that time equivalent to about 1 USD) between a private pot and a common pot. The former represents the participants' personal pot – it is up to him/her as to decide what to do with these funds, while the latter represents a communal pot – it is up to the club, as a team, to decide what is to be done with these funds. In order to guarantee confidentiality while making this decision, we recorded the individual split decisions in private, and ourselves also contributed to the common pot. In this way, no-one could effectively deduce how much others had contributed to the common pot.

The final twist of the game is the following: the money put in the common pot would be multiplied by 2, while the money put in the private pot remains as is. This implies that, if decision-making between the various club members were perfectly cooperative, all members should contribute the full amount of 400 Kw in the common pot. (Protocol is available on request from the researchers).

Figure 15 reports the results. One can see that most farmers (95%) contribute a non-zero amount to the common pot. The median contribution is 150, the mean is 170 and the standard deviation is 120.

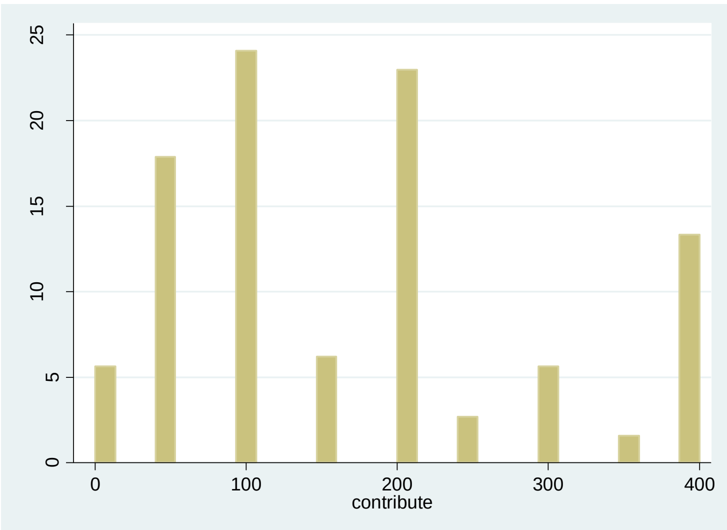

Figure 15: Distribution of contribution to the common pot [in Kw]

4. Baseline differences between treatment and control

In this section, we test for baseline differences between the treatment and control group.

We start by presenting the results of t-tests, testing for baseline differences in village characteristics between treatment and control villages, and move on to – using a basic regression approach – testing for differences between baseline differences in household characteristics in these villages.

As the sample in treatment villages consists of both club-members as well as non-club members, a status which was self-selected by the individuals, we further also test for the (observable) differences between club and non-club members in the treatment villages using a regression approach.

We conclude this section by taking a look at the treatment villages selected for demonstration plots – and check whether these villages – as we would expect them to be – are different from other villages in the treatment group. We use a series of t-test to do so.

4.1. Village characteristics

Table 37 reports the results of a series of t-tests with unequal variances. We note that there is a substantial amount of variance within both treatment and control villages. Comparing the averages between the treatment and control villages, one can safely say that these are fairly close to one another across all variables selected below. As a result, we cannot detect any significant differences between treatment and control villages.

Table 37: Differences between treatment and control villages at baseline

Variable	Mean (St. dev) Control	Mean (St. dev) Treatment	Difference in means (Treatment – Control)	P-value
Number of households	78 (88)	67 (61)	-10	0.26
Distance to a paved road (km)	2.71 (5.48)	2.26 (4.40)	-0.45	0.48
Distance to a national highway (km)	7.25 (9.14)	8.17 (10.95)	0.91	0.47
% Area under soy cultivation 2014	25.78 (21.19)	22.44 (14.88)	-3.33	0.15
% Farmer cultivating soy 2014	40.04 (27.77)	41.60 (26.44)	1.12	0.74
Market for pesticides, fertilizer and seeds all year	13.44 (10.69)	13.57 (10.26)	0.13	0.91
Market for pesticides, fertilizer and seeds during season	6.18 (6.53)	5.72 (6.10)	-1.46	0.56
Bank or formal credit organization	15.70 (10.91)	15.87 (12.01)	0.16	0.91
Distance to closest market to sell produce	5.76 (9.29)	4.28 (6.59)	-1.48	0.15
Market for fresh fruits and vegetables	2.83 (5.40)	2.71 (3.38)	-0.12	0.82
# of Visits by Government	6.11 (16.12)	6.71 (13.34)	0.60	0.75
# of Visits by NGO's	15.48 (26.44)	17.89 (29.48)	2.45	0.49

4.2. Household characteristics

Table 38 reports the result of the following regression specification, where $y_{i,j}$ is the dependent variable of interest and T_j represents the village level treatment. To take into account that the treatment was administered at the village level, we cluster the standard error at the village level. In Table 38, we report the regression estimate of β_1, together with the standard error and P-value.

$$y_{i,j} = \beta_0 + \beta_1 T_j + \mu_j + \varepsilon_{i,j} \quad (1)$$

We find no significant differences between baseline characteristics of treatment and control households for most of the variables, including the age of the household head, the acreage of owned land, and the education of the household head. The treatment households do have a higher chance to be headed by a woman and tend to have more household members.

Importantly, we find no significant differences between most of our outcome variables of the evaluation: agricultural yields of soybean and groundnut, adoption (in particular the acreage under soy cultivation) and perceived per-acre yields of maize, groundnut and soybean. However, we do note a significant difference between treatment and control households at baseline in terms of maize yield: surprisingly, treatment households have lower maize yields at baseline as well as the percentage of maize marketed – treatment households market less of their maize harvest. Treatment households also tend to have slightly lower overall per-capita income.

Table 38: Differences between treatment and control households at baseline

Variable	Regression Coefficient Estimate	Standard Error Estimate	P-value
Characteristics			
Age household head (years)	-0.11	0.62	0.865
Number of household members	0.23	0.12	0.048
Gender household head (1=female, 0=male)	0.03	0.02	0.052
Acreage of owned land (acre)	-0.27	0.25	0.269
Asset wealth (Mk)	216,882	412,586	0.600
Education household head (years)	0.21	0.18	0.244
Outcome variables			
Acreage under soy cultivation (acre) in 2013-14	0.06	0.07	0.338
Soy yield (per acre) – 50kg bags shelled	-0.30	0.44	0.498
Maize yield (per ace) – 50kg bags shelled	-1.61	0.72	0.025
Groundnut yields (per acre) – 50kg bags unshelled	0.08	0.57	0.885
Maize marketed or intend to market in 2013-14 as percentage of the harvest	-0.04	0.01	0.007
Per capita agricultural income (Kw)	-2,011	1,223	0.101
Maize, perceived yield (in 50 kg per acre, shelled)	0.03	0.84	0.969
Groundnut, perceived yield (in 50 kg per acre, unshelled)	0.50	0.68	0.464
Soybean, perceived yield (in 50 kg per acre, shelled)	-0.11	0.35	0.753

Note: Yield per acre is reported for mono-cropped fields only. Outliers above the top and bottom 1 percentile are dropped for the yield calculations. For income per capita outliers above the 95th and below the 5th percentile are dropped.

4.3. Club and non-club members

In Table 39, we investigate the baseline differences between club and non-club members, using the regression specification (2). Again, we cluster standard errors at the village level. Note that as the control villages do not have any clubs formed in the first place, we limit the sample to treatment villages which have clubs present only.

$$y_{i,j} = \beta_0 + \beta_1 C_i + \mu_j + \varepsilon_{i,j} (2)$$

We note significant differences between club member and non-club members in this context. In particular, club members are, on average, more educated, have a higher per-capita income and have more optimistic beliefs in terms of perceived yields. On the other hand, they also tend to own less land and harvested lower yields of soybean in 2013-14.

Without any further analysis, we are cautious not to over-interpret these results. It is possible that farmers who had a previous bad experience in soy farming, but in general are well established farmers, were exactly the ones who were attracted by the CDI program.

Table 39: Differences between households with and without club members at baseline

Variable	Regression Coefficient Estimate	Standard Error Estimate	P-value
Characteristics			
Age household head (years)	0.42	1.05	0.691
Number of household members	0.28	0.19	0.132
Gender household head (1=female, 0=male)	-0.02	0.02	0.486
Acreage of owned land (acre)	-1.49	0.75	0.061
Asset wealth (Mk)	117,387	123,384	0.344
Education household head (years)	0.73	0.24	0.004
Outcome Variable			
Acreage under soy cultivation (acre) in 2013-14	0.07	0.06	0.236
Soy yield (per acre) – 50 kg bags shelled	-1.41	0.59	0.022
Maize yield (per ace) – 50 kg bags shelled	1.74	1.33	0.195
Groundnut yields (per acre)–50 kg bags unshelled	0.94	0.79	0.234
Maize marketed or intend to market in 2013-14 as percentage of the harvest	0.15	0.14	0.261
Per capita agricultural income (Kw)	6,126	1,922	0.002
Maize, perceived yield (in 50 kg per acre, shelled)	1.95	1.24	0.121
Groundnut, perceived yield (in 50 kg per acre, unshelled)	3.69	0.93	0.000
Soybean, perceived yield (in 50 kg per acre, shelled)	1.42	0.49	0.005

Note: Yield per acre is reported for mono-cropped fields only. Outliers above the top and bottom 1 percentile are dropped for the yield calculations. For income per capita outliers above the 95th and below the 5th percentile are dropped.

4.4. Demonstration plot and non-demonstration plot villages

Finally, Table 40 reports the results of a series of t-tests with unequal variances comparing the selected demonstration villages, with villages in the treatment group not selected as demonstration plot sites. Note that the results of these particular tests need to be taking with a grain of salt, as we only have 19 demonstration plot villages (N=19 of the demonstration plot group).

Despite the small sample size of the demonstration plot villages, we find some notable differences between these villages and the other treatment villages. In particular, while not statistically significant at the 10% level, demonstration plot villages appear to have more experience in cultivating soy (at baseline) and tend to be closer to markets. This would confirm the strategic selection of CDI of these demonstration plot villages.

Table 40: Differences between demonstration plot and non-demonstration plot villages at baseline

Variable	Mean (St. dev) Non-demo plot	Mean (St. dev) Demo plot	Difference in means (non demo-plot – demo plot)	P-value
Number of households	74 (78)	52 (8)	22	0.028
Distance to a paved road (km)	2.44 (5.03)	3.35 (2.94)	-0.90	0.302
Distance to a national highway (km)	7.68 (10.20)	6.64 (6.13)	1.03	0.564
% Area under soy cultivation 2014	24 (18)	28 (19)	-4.06	0.466
% Farmer cultivating soy 2014	40 (26)	52 (26)	-11	0.1339
Market for pesticides, fertilizer and seeds all year	13.57 (10.54)	11.64 (7.01)	1.93	0.346
Market for pesticides, fertilizer and seeds during season	5.97 (6.41)	4.68 (2.79)	1.29	0.145
Bank or formal credit organization	15.76 (11.53)	15.42 (8.53)	0.33	0.891
Distance to closest market to sell produce	5.12 (8.16)	2.85 (4.75)	2.27	0.11
Market for fresh fruits and vegetables	2.75 (4.55)	2.55 (2.43)	0.19	0.78
# of Visits by Government	6.43 (15.00)	4.92 (5.94)	1.50	0.426
# of Visits by NGO's	16.92 (28.27)	11.25 (15.01)	5.67	0.246

5. Revised power computations

Power measures the probability that we avoid Type II errors, or failure to reject the null hypothesis when it is in fact false (false negative). We conducted the power analysis using an 80% power using data from the baseline survey.

We follow Glennester and Takavarasha (2013) to compute the Minimum Detectable Effect Size (MDE) using Stata's *sampsi* and *sampclus* commands given a sample size of 500 clustered in groups of 10. This implies that we focus on the midline evaluation, as we effectively assume zero attrition. We note that our results are not very sensitive to this assumption, and some attrition would not change the computations much.

We focus on the following outcome variables: yields of maize, soybean and groundnut, acreage of soybean (as a measure of adoption), per-capita rainy season agricultural income, and beliefs regarding maize, soybean and groundnut. At midline, we will in addition look at a knowledge index as an outcome variable. We have no knowledge index at baseline, but we propose to follow Kondyliis and Mueller 2012 in a study of demonstration plot effectiveness in Mozambique. They report a 24% increase in knowledge about sustainable land management; using their knowledge variable, an index with a mean of 0.24 and standard deviation of 0.11, and assuming a rho of 0.2; our MDE would be 16%.

In our computations of mean, standard deviation and within-cluster correlation (using Stata's *loneway* command), we do not omit any outliers, i.e., we consider the full distribution.

Table 41 reports the results. These MDEs are the differences we can pick up comparing between the control and treatment group at midline. We note that the MDE are around 20% of the yields, and around 10-15% for the perceived yields. The MDE, as expected, is fairly large for the agricultural income. Hence, we expect our prospects to detect any effect on income to be limited at baseline. At endline, we can trim the income distribution and drop outliers, as well as use control variables to increase precisions of the effect estimates.

Note that just 18% of households are headed by a female. Hence, we do not expect to be able to pick up much of the gender-specific effects, unless the effects sizes between male and female-headed households are substantial.

The still somewhat large MDEs imply a discussion on general equilibrium effects. In particular, one could imagine the AFM project affecting the local economy through the price/income channel: the increase in maize and soy yields could affect the price of maize and soy in the local market through the technology treadmill effect. The provision of free inputs for the demonstration plot – as they are limited to demo plots only – are unlikely to affect the local pricing market directly, but the increased demand for ISFM related inputs might affect the price of soy seeds, for instance. The increased purchasing power of farmers in the region might have a multiplier effect on the local economy as well. Using the Malawi Household Survey Dataset – a repeated cross-section conducted every 4 to 5 year (rounds of 2004-05, and 2010-11 and later rounds), we can, at the endline stage, compare the average prices in the areas in which CDI is working with the areas in which CDI is not working using a difference-in-difference approach.

Table 41: MDE of household level outcome variables

Outcome variable	*Results*	*Assumptions*			*Unit*
	MDE	*Mean*	*St. Dev.*	*Rho*	
Soy yield (per acre)	20%	2.5	2.2	0.084	50 kg bags shelled
Maize yield (per acre)	20%	12.4	10.2	0.115	50 kg bags shelled
Groundnut yield (per acre)	20%	10.3	8.2	0.082	50 kg bags unshelled
Area under soy cultivation in 2013-14 (acre)	23%	0.77	0.84	0.025	Acre
Per-capita agricultural income (Kw)	63%	30,566	86,214	0.053	Kw
Maize, perceived yield (per acre)	12%	28.1	15.5	0.083	50 kg bags shelled
Soybean, perceived yield (per acre)	11%	12.1	7.1	0.045	50 kg bags shelled
Groundnut, perceived yield	15%	19.9	13.2	0.070	50 kg bags unshelled

Data Appendix

Table A1: Relation of individual member to respondent

	Number of individuals	Percentage
Respondent	2,509	19.53
Wife/husband	2,058	16.02
Child	6,806	52.96
Brother/sister	206	1.60
Parent	104	0.81
Aunt/uncle	9	0.07
Grand parent	25	0.19
Other relative	1,065	8.29
Unrelated	68	0.53
TOTAL	12,850	100.00