Midline Report

Study Title: The effect of demonstration plots and the warehouse receipt system on ISFM adoption, yield and income of smallholder farmers: a study from Malawi's Anchor Farms

Thematic Window 4 TW4.1018

Investigators: Ephraim Chirwa, Hope Michelson, Chris Barrett, Cheryl Palm, Annemie Maertens and Wezi Mhango

Persons prepared the report: Annemie Maertens, Eric Kaima, Hope Michelson and Wezi Mhango

Date of the report: 30 April 2016

Table of 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. In 2014-15, we randomized 250 villages into (i) a control group, and (ii) a treatment group. The villages in the treatment group were invited to form farmer clubs, elect a lead farmer and participate in CDI's activities under (i) above – production knowledge dissemination. In this report, we present the first results of this exercise, assessing the effects of the CDI program on farmers' adoption behavior, knowledge and beliefs, prices and profits. In the next 3 years of the project, we aim to evaluate also the effects of CDI's market interventions, as well as establish the longer term effects of (i). In addition to identifying the treatment effects, the detailed panel data collected will permit us to 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.

2. Sample at midline

2.1. Villages

Our overall sample at baseline consistent of 250 villages: 125 were selected into treatment and 125 into control. The midline sample consists of 100 out of the 250 villages (see Table 1): 54 in Dowa district and 46 in Kasungu district. Of these, we selected 90 at random (stratified per treatment/control status and district location) and 10 purposefully. The goal of the purposive sample was to include all 17 demonstration plot villages in the sample which will allow us to do a detailed analysis of information flows and learning within these 17 demonstration plot villages.[1] Of these 17 demonstration plot villages, 7 are included in the random sample, while 10 are not.

2.2. Households

Recall that the CDI program works through farmer clubs. Hence, at the start of the program, all treatment villages were invited to a sensitization meeting and asked to form club. About 70% of treatment villages did so. As we expected the effects of the treatment to depend on whether or not an individual is member of a club, and random sampling would under-sample club-members and most likely not include the leader of the club, we stratified the sample in the treatment villages (with clubs) between club members and non-club members. In particular: we selected 5 club members (including the leader) and 5 non-club members (randomly from both census listings). In the control villages and treatment villages without clubs we just selected 10 households randomly from a census list.

As CDI club membership was self-selected, we could expect CDI club members to be different from non-club members, both along observable and non-observable characteristics. This implies that, unlike in the control villages where no self-selection into CDI clubs happened[2], the sample

[1] This selection mechanism of the midline villages differs from what we had originally proposed. In the previous progress report and the midline analysis plan, we had proposed to pre-select 50 villages in a non-random manner: all villages which had demonstration plots and all villages which have soil samples. This would leave us with a random sample of just 50 villages. The goal of this design was to maximize the number of villages for which we also had soil sampling information. However, concerns about a small treatment effect lead us to alter this strategy and select 90 villages at random (half treatment, half control) and 10 purposefully (additional demonstration plot villages). In this way, we have 90 villages to do the midline evaluation with. To deal now with the lack of soil sample information (only 25 villages in our midline sample also have soil samples collected in them), we propose to exploit the links between subjective and objective soil sample information, in particular we intend to create a mapping between the perceived soil quality and the chemical/structural characteristics of the sample. Recall also that for the endline, we will return to all 250 villages, so this will no longer be an issue then.

[2] Unfortunately, logistical limitations did not allow CDI to form clubs in the control villages. This is because the formation of clubs coincided with a first information session about the CDI program, and hence the treatment and club-formation are intrinsically linked to each other. If this were not the case, one

drawn in the treatment villages is unlikely to be representative of the village as a whole. We will take this issue into account in the analysis using a propensity-score-matching inspired method.

2.3. Attrition from baseline

The midline attrition rate is 5% - which is in line with our assumptions in the project proposal. Specifically, there were 51 households who were present at the baseline who were not present during the midline survey. The households who left the sample are well distributed geographically; they come from both districts and from both treatment and control groups. The households who left the sample have household heads who are slightly younger and (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 [These results are available on request]. Interestingly, once we control for observable household and head characteristics, CDI club members are less likely to leave the sample than non CDI club members. This is an issue which will require further investigation.

3. Participation into the CDI program

As noted in the baseline report, the CDI program consists of two primary activities in the 2014-15 season: demonstration plots and farmer field days. The demonstration plots were established in 17 villages selected strategically by CDI and managed by CDI clubs (guided by a CDI extension agent) throughout the season.[3] Two farmer field days were held in two locations at the end of the 2014-15 growing season and all CDI club members were invited to attend.

All activities centered around the farmer clubs: all treatment villages were invited to form farmer clubs, and CDI activities took place at the club level (i.e., the demonstration plots were managed by the club). As the clubs were self-selected, it is of utmost importance to understand which villages formed clubs, and among the villages that did form clubs, which farmers joined the clubs. We address these issues in Sections 3.1 and Section 3.2, respectively. In Section 3.3 we present the descriptive statistics of the participation module in the midline – i.e. documenting participation in the various CDI activities that took place in the 2014-15 season.

3.1. Which type of village forms a club?

All villages in the treatment group were invited by the CDI extension officer (whom in their turn asked the government extension agents to invite the relevant village heads etc.) to attend CDI's sensitization meetings in the area. It was at these meetings that the CDI project was introduced and the villages were asked to form clubs. We hypothese that whether or not a village formed a club would depend on a variety of factors. Distance to sensitization meeting location and prior relationships with the government extension agents might affect whether or not anyone from the village attends the sensitization meeting (see the baseline report on descriptive statistics on

---

could simply compare club members in treatment villages to club members in control villages; as well as non-club members in treatment villages and non-club members in control villages.

[3] Recall that one demonstration plot was established in a control village.

sensitization meeting attendance). Village characteristics such as acreage under soy cultivation, ethnic composition, the existence of village organizations might affect the village's interest to participate in the program and ability to come together in a group.

In Table 2 we report the results of a Probit regression using baseline village characteristics to predict whether or not a village had formed a CDI club as of October 2016. We consider village level characteristics at baseline collected in the baseline village questionnaire and how they map up into the likelihood of the village forming a club. Note that this sample includes the 125 treatment villages only. Note also that we do not consider household level characteristics as the household sample in villages with and without clubs are not comparable due to stratification issues (see earlier discussion). We hypothesized that the following factors might play a role in whether a village formed a farmer club: (i) distance to roads, the main highway and agricultural services, (ii) previous contact with extension agents, (iv) village size, (v) ethnic composition of the village, and (vi) past organizational experience. The results show that villages with fewer households are more likely to form a CDI club, as are villages with fewer males, with a larger number of village organizations and more members in their organization, on average. Perhaps somewhat surprisingly, an increase in the number of ethnic groups in a village does not decrease the likelihood of forming a CDI club. In future extensions, we will include distance to the sensitization meeting (using GPS data) and the results of the public goods game as regressors. In addition, we aim to shed more light on group composition as well – conditional on village characteristics. For instance, informal observations suggest that villages which are close to markets tend to have more women in their clubs, possibly indicating a high opportunity cost of time for the men.

We use results from the regression presented in Table 2 to predict out-of-sample whether or not a village in the control group is likely to form a club. When this predicted probability exceeds 0.5, we classify this control village as one that is likely to form a club. Using this method we classify 80% of the control villages as villages which are likely to have clubs. Note that this percentage is similar in the treatment villages. Using the treatment sample only, we can test the accurateness of this prediction: We find that we do fairly well: we classify 75% of the treatment villages in the correct status (Table 3 reports the results).

3.2. Which households participate in CDI farmer clubs?

In Table 4, we present a Probit regression analyzing the relationship between household (specifically, any member of the household) CDI club participation and a vector of baseline household characteristics. We only consider only the treatment villages with a CDI club in this analysis. We find that households with more members and with higher levels of education are more likely to join a CDI club. Even though the coefficient is not statistically significant at the 10% level – the data seems to suggest that women-headed households might also be more likely to be part of CDI.

We use this regression to predict individual CDI club membership among the control villages (restricting ourselves to the control villages who we predicted – in the previous section – to have clubs). Again, we test the accuracy of this prediction on the treatment villages with clubs: 60% of the household's status is correctly predicted.

In future analysis, we plan to extend the regression in Table 4 to include also non-linear effects. In particular, one could imagine a model in which households weigh the costs and benefits of participation. Households with substantial assets would also have a higher opportunity costs of time, as might households with very few assets, leaving only the households with a medium level of asset holdings to participate. These can give rise to non-linearities in the participation decision. This might especially be the case if the underlying asset distribution is non-normal as well. In Figure 1 we plot the histogram of the total value of assets (in log Malawian Kwacha). The bimodal distribution suggests possible evidence of a poverty trap, consistent our hypothesis of non-linearities.

3.3. Participation in the CDI activities

The CDI activities are diverse. CDI started of their activities in the project area with a sensitization meeting in September of 2014 (which all farmers from treatment villages were requested to attend) and followed up with a lead farmer training in December 2014, inviting the lead farmers of the clubs in the treatment villages. Then, later in December 2014, CDI established 19 demonstration plots in 17 strategically selected villages and later on in the season in December 2014 – at the time of harvest in March-April 2015 – invited club members from all the treatment villages to attend a field day.

In this section, we focus on understanding participation in the field day and demonstration plot activities. See the baseline report for statistics on participation in the sensitization meetings and lead farmer training.

3.3.1. Knowledge about CDI

Table 5 presents information about farmers' knowledge about CDI. In the treatment villages 81% of respondents had heard about CDI or the AFP. In the control villages 41% had heard about CDI or the AFP. Table 6 shows knowledge about CDI's local extension agents working on the AFP. Merely 30% of respondents in the treatment villages know the CDI agent by name, and very few – only 0.03% of the control village respondents knows the CDI agent by name, suggesting very limited spillovers. In the sample as a whole, 46 respondents reported that the CDI agent had visited their farm or house (all of these respondents live in treatment villages).

3.3.2. Field-days

Field days for the 2014-2015 growing season were conducted under the theme: *Creating wealth through modern legume production"*, and had the following objectives: (A) to increase awareness of the general public about the CDI Anchor Farm, and (B) to disseminate productivity enhancing technologies and practices to farmers, extension workers and partners, especially, best practice soy bean agronomy and Integrated Soil Fertility Management (ISFM) with Conservation Agriculture with fertilizer trees; and (C) to provide a platform to input suppliers and service providers to interface directly with farmers regarding products and services they provide. CDI organized two field days during the mid-season for farmers from the two study areas.

The first field day was held on 19 March 2015 at Kapinya village of Mtunthama EPA and all clubs from the treatment villages in the area were invited to attend. This village was chosen because it had the most successful maize and soybean demonstration plot among all 11 demonstration plots implemented in that EPA during the season. No such field day took place in Chibvala (as most crops on the demonstration plots did not perform well due to the dry spell). Instead, a few farmers from each club in Chibvala were invited to a field day at Galeka village in Lisasadzi EPA in Kasungu District on 30 March 2015. A total number of 606 farmers attended the field day in Mtunthama EPA and 72 farmers from Chibvala EPA attended the field day at Galeka Village in Lisasadzi EPA. Lead farmers and other members from the club hosting the field day took lead in demonstrating to visitors how they went about in applying new technologies in soybean and maize plots. In addition to learning new technologies, farmers were shown how to prepare various food items from soy beans. In addition, a certified agro- dealer known as Farmers` Hub was contracted by CDI to sell agriculture equipment and chemicals to farmers in the project areas in the 2014-15 growing season. They had a display of their products that included, among others; inoculant, herbicides, insecticides, sprayers, maize-shellers, and vegetable seed (cabbage, tomato, onion) and explained to farmers how to use and where to buy these products.

Table 7 reports the number of respondents who had heard about one of the local CDI field-days: 40% of the respondents in treatment villages had heard about the CDI field-days versus 17% of respondents in the control villages. Table 8 reports on the number of respondents who had at least one household member who participated in the CDI field-days. 14% of households in the treatment villages participated in the CDI field-days versus 3% of households in the control villages. It is notable that the majority of this participation took place in Kasungu, and that participation was not restricted to individuals from the 17 demonstration plot villages alone: 12 out of 29 treatment villages without demonstration plots had one or more participant in the CDI field-days. Figure 2 shows the number of farmers (per village) that participate in the CDI field-days in the treatment villages with clubs. In about 65% of the villages at least one person within our sample attended the field-day (note that the other 35% might have participants, but these are not member of our sample). Most villages with at least one participant would have one, two, or three members participating. Keeping in mind the structure of the sample – in which 5 club members were sampled – this implies that rarely all club members would attend the field-day. Table 9 reports on the perceived usefulness of the field days among participants: 83% of the participants reported them to be "very useful".

In future analysis, we will investigate how the individuals that attended the CDI field-days have shared information with the other residents of their villages. The dynamics of information sharing in agricultural extension is crucial to understand given the numerous projects that follow a model like CDI's implicitly depend on information trickling down from one network node to the other. In this case, even if just 14% of individuals in our sample attend the CDI field-days – successful information spreading might still be possible and would depend on the way people learn and share information.

3.3.3. Demonstration plot activities

Recall that CDI established demonstration plots in 17 villages in the 2014-15 season. In the midline questionnaire, we asked all respondents about their knowledge and participation in the activities that took place on these demonstration plots.[4]

Using the midline data, we find that clubs who managed a demonstration plot met, on average, seven times. Overall, most meetings were well attended (as also evidenced in the agronomic report which reports the CDI participation data we collected – which represent census data while this report is based on a sample). Again, around 70% of attendees considered these meetings very useful, and another 25% considered them useful.

While the focus group report confirms that clubs who manage demonstration plots are very active throughout the season, this is not necessarily true for the clubs who were not assigned demonstration plot. We are currently working together with a PhD student to collect club-level data among all CDI's clubs. Using these data we can continue to follow up the club dynamics among all clubs – post baseline.

In addition, knowledge of the activities on the demonstration plots – with exception of the field-days – appear to be largely constrained to the club-members within that village only. Only half of non-CDI club members in CDI demonstration plot villages appeared to be aware of the presence of a CDI demonstration plot in their village. Outside the demonstration plot villages, only 16% of respondents in the treatment group stated to be aware of any CDI demonstration plot in another village. These statistics are confirmed in the agronomic report where we report attendance for each demonstration plot activity: few people from outside of the village attended the various demonstration plot activities.

4. Analysis and results

4.1. Conceptual framework

Basic specification:

Denote the outcome variable of interest of individual i in village j as as y_{ij}. Assuming that treatment and control households are similar at baseline, we first compute the IIT (Intention-to-Treat) estimate of the **average treatment** effects as follows:

$$y_{ij} = \alpha + \beta_1 T_j + \mu_j + \varepsilon_{ij} \quad \text{(1A)}$$

[4] An interesting feature came to our attention during the midline focus groups: certain CDI clubs that were not assigned a demonstration plot by CDI took the initiative to start their own demonstration plot. Table 10 tabulates the answer to the question: 'Does your club manage a demonstration plot' for all CDI club members: we find that 58 respondents who answer yes to this question, even though they were not in one of the 17 demonstration plot villages (interestingly, there are seven CDI club members who live in demonstration plot villages who claim not to be involved and possibly misunderstood the question). We intend to explore this spill-over phenomenon further in future analysis.

Where $T_j \in \{0,1\}$ indicates whether the individual belongs to a treatment village or not[5], μ_j is a village level error term while ε_{ij} is an individual level error term. As correlation between observations can result in incorrect standard error estimates, we cluster standard errors at the village level to allow for arbitrary correlation between village members which affects y_{ij}.[6]

To **increase the precision** of the β_1 estimate, we, in selected regressions, also include baseline household characteristics in specification (1) by using (2):

$$y_{ij} = \alpha + \beta_1 T_j + \gamma \bar{X}_{ij} + \mu_j + \varepsilon_{ij} \quad (2)$$

Where $\bar{X}_{ij}$ denotes a vector of individual level characteristics at baseline: gender of the household head, acreage owned, education of the household head, number of adult household members, total value of the assets (in MK), time preferences and soil quality (measured as the average rental value of each acre of land in 1000 MK).

In selected regression, we explore **heterogeneity in treatment effects** as a function of baseline household characteristics to better understand the mechanisms driving our results and how likely they are to generalize to other contexts. In particular, we look at the role of baseline soil quality and baseline yield beliefs. This yields specification (3):

$$y_{ij} = \alpha + \beta_1 T_j + \gamma \bar{X}_{ij} + \beta_2 T_j \bar{X}_{ij} + \mu_j + \varepsilon_{ij} \quad (3)$$

Specification taking into account self-selection into clubs:

The method outlined next is inspired by the method of Propensity Score Matching. For the treatment villages with clubs, we use the baseline observable characteristics to predict self-selection into a CDI through a Probit specification as in specification (4):

$$P\left(club_{ij,treatment} = 1\right) = \alpha + \beta_1 \bar{X}_{ij} + \mu_j + \varepsilon_{ij} \quad (4)$$

Using (4), we predict CDI club membership in the control villages (which we predicted to have clubs in the first place) as outlined in specification (5), where $\hat{\alpha}$ and $\hat{\beta}_1$ of (5) are the estimates obtained from regression specification (4).

$$P\left(club_{ij,control} = 1\right) = \hat{\alpha} + \widehat{\beta_1} \bar{X}_{ij} \quad (5)$$

We then proceed to set the cut-off rule at a probability of 0.5. If the probability is larger than 0.5, we assume club-membership, while if it is smaller than 0.5 we would assume non-membership. Using this cut-off rule, we divide the farmers in the control sample (in villages which are

[5] In treatment villages, lead farmers were invited to attend a course in December 2014, and all farmers in the village farmer clubs were invited to attend demonstration on nearby demonstration plot sites throughout the season.

[6] As the numbers of clusters in the midline is on the small side, we intend – in future analysis - to compare this standard error specification with a village–fixed effect with robust standard error specification. For a discussion, see also Cameron and Miller (2013): A Practitioner's Guide to Cluster-Robust Inference.

predicted to have clubs) into two groups: the group of the "would-be" club members, and a complementary group of not "would-be" club members.

We can then proceed and estimate the effect of the treatment for CDI club members and CDI non-club members through two separate specifications as in (6A) and (6B), i.e., we effectively run these two regressions for the two mutually exclusive sub-samples constructed: (i) the sub-sample of club-members in the treatment villages and predicted club members in the control villages; and (ii) the sub-sample of non-club members and predicted non-club members in the control villages:

$$y_{ij,club} = \alpha + \beta_{1,\text{club}} T_j + \mu_j + \varepsilon_{ij} \text{ (6A)}$$

$$y_{ij,non,club} = \alpha + \beta_{1,\text{non-club}} T_j + \mu_j + \varepsilon_{ij} \text{(6B)}$$

In addition, we use a similar method to predict whether or not a village forms a CDI club, and compare the households in the treatment villages without clubs with the households in the control villages which we predict would not form any club either using specification (6C):

$$y_{ij,non,club;non-club\ village} = \alpha + \beta_{1,\text{non-club;non-club village}} T_j + \mu_j + \varepsilon_{ij} \text{(6C)}$$

Sample:

We use the sample of 90 villages which have been randomly selected. As the number of outcomes is very large, we focus on the outcomes presented in the midline analysis plan, in addition to some selected intermediate indicators. As discussed above, we present the results for three groups of farmers:

(i) Farmers in treatment villages without CDI clubs (with as comparison group the control variables in villages which we predict would not form a CDI club either when given the opportunity)
(ii) Farmers in treatment villages with a CDI club but who are not members of the CDI club (with again a comparison group constructed from the control sample)
(iii) Farmers in the treatment villages with a CDI club who are members of the CDI club (with again a comparison group constructed from the control sample)

To limit the total number of results reported in this report, we focus on average intent-to-treat effect (the treatment effects on the treated results using an instrumental variable specification where we instrument treatment for participation in the field days are available on request[7]), and

[7] There are two additional reasons why we currently refrain from reporting the treatment-effect-on-the treated. First, even if a farmer did not attend a field-day, he/she is likely to hear about it from their network. Hence, a complete analysis of the treatment-effect-on-the-treated would need to incorporate these spillovers. We intend to address these spillovers in future analysis. Second, one needs to give the demonstration plot villages a special status as CDI members there are likely to learn also from the demonstration plot activities. Given that these villages were purposefully selected, we propose combined

only report the regression results with and without control variables in selected cases – again guided by the midline analysis plan.

4.2. Descriptive statistics on (selected) dependent variables

In this section, we focus on the outcome variables which are unique to the midline: (i) Knowledge about ISFM practices, (ii) Beliefs about yields, and (iii) planned adoption of ISFM practices. For the outcome variables which were also covered in the baseline: (iv) Acreage, (v) Prices, (vi) Yields – see the baseline report.

The midline questionnaire incorporated a set of 20 questions which aimed to test knowledge on ISFM adoption. We developed this set of questions jointly with CDI. [See the midline questionnaire on the exact questions] We code the answers as correct/incorrect and compute a total knowledge score [out of 20]. Figure 3 presents the histogram of this knowledge score. The average score is 7 with a standard deviation of 2. In Table 11, we present the results of the knowledge test, disaggregated by questions. It is notable that most respondents do have a general sense of ISFM practices but often lack detailed knowledge about the type of chemicals involved and in particular how about the quantity to use.

In the planned adoption module, we asked the respondent about their intention to adopt selected ISFM practices for the 2015-16 season. We first asked about the four main CDI crops: Soybean, Maize, Groundnut and Beans. Then we followed up with questions regarding the variety, whether or not they plan to use inoculation (in the case of soybean) and how long they have been growing this crop. The data reveals that most farmers plan to cultivate the CDI crops – over 90% for all crops and almost 100% in the case of maize (not that surprising in the case of Malawi). When asked about their experience with these crops, very few (less than 5%), state that this upcoming season would be the first time they had tried to cultivate the crop. Hence, we do not expect much effect of the CDI program on this margin. Instead: we propose to focus on crop practices. First, we inquire about the variety. As CDI recommends specific varieties, this might be an indicator of the effect of the program. The recommended variety for soybean is Squire. The recommended variety for groundnut is CG7. The recommended variety for maize is SC719. The recommended variety for common bean is Kholopete. Table 12 reports the results. It is interesting to note that very few farmers intend to adopt the recommended CDI soybean variety: Squire. Finally, Table 13 present the inoculation plans: despite being quite aware of the benefits of the technology (see Table 11), few farmers intend to adopt the practice: 13% (among those that plan to adopt soy).

In the next section of the planned adoption module, we ask the farmer about a set of ISFM practices that CDI promotes: mixed cropping/intercropping, crop rotation, the use of organic and inorganic fertilizer, the use of herbicide, the use of pesticide, the use of fungicide and the planting of fertilizer trees. Table 14 reports the results. Most farmers plan to adopt mixed cropping and rotation practices, as well as organic and inorganic fertilizers. Few farmers plan to use fungicides, herbicides, pesticides or fertilizer trees. When asked about the reason for the

panel data / propensity score matching approach. As the method requires us to limit to outcome variables which are included in at least two surveys, we propose to postpone the final analysis to the endline (even though we intend to do a preliminary propensity score approach using midline data only).

non-adoption, farmers cite issues with market access, credit constraints and prices. Interestingly, when asked about previous experience with these technologies, farmers who do plan to adopt fungicide, herbicides, pesticides and fertilizer trees note that this upcoming season will be their first season, possibly indicating an effect of CDI activities.

The midline questionnaire also included an extended beliefs module, inquiring about the farmer's subjective beliefs regarding the yield of local maize, hybrid maize, soybean and groundnut under a range of different weather conditions. Figure 4 reports the results. Farmers, on average, expect a yield of 12 50kg bags of soybeans (shelled) for one acre of land, 24 50kg bags/acre of hybrid maize (shelled), and just 15 50kg bags/acre of local maize (shelled) and 18 50 kg bags/acre of groundnut (unshelled, dried). Interestingly, when asked about whether they expect a better/worse yield 3 years from now, 60% of respondents mention a higher number for both hybrid maize and soy, indicating that they might expect longer term effects of the ISFM practices they might intend to use (or other technologies they intend to adopt). In future analysis, we plan to use the full distribution of these beliefs to shed light on the learning process, in particular in the demonstration plot villages. For instance: Do farmers become less uncertain about the performance of these crops over time as they observe demonstration plot performance? Do farmers with 'worse' soils have more uncertain expectations and are they less likely to learn anything useful from the demonstration plots (as these plots are often on the 'better' soils)?

4.3. Results

This section follows the structure of the midline analysis plan. We first present the results on outcome variables pertaining to the 2014-15 season, and then present some selected results on intermediate outcome variables: knowledge, beliefs and planned adoption.

Table 15 presents the average treatment effect of the CDI program on the likelihood that the farmer uses inorganic fertilizer in the 2014-15 season in columns (1), (3) and (5) (we use a limited probability model OLS regression; results of a probit regression are similar and available on request). We find that the CDI program significantly increases the likelihood of using inorganic fertilizer among non-club farmers who live in treatment villages with a CDI club. The effect size is economically significant, estimated at 26 percentage points. In Columns (2), (4) and (6), we present the treatment effects, this time controlling for gender of the household head, acreage owned, education of the household head, number of adult household members, total value of the assets (in MK), time preferences and soil quality (measured as the average rental value of each acre of land in 1000 MK). We find that increased soil quality decreases the likelihood of using inorganic fertilizer – but only for farmers living in villages where no CDI club was formed and report a positive interaction effect of soil quality with the treatment. Note however that this measure of soil quality is rather crude and we will need to conduct further analysis to narrow this down to the relevant dimensions of soil quality using both the soils sampling results as well as the perceived soil quality results. In addition, this regression would benefit from a plot-level approach – as one would expect this decision to be made at the plot level and the household level. We intend to make these extensions in the coming year.

Table 16 presents the average treatment effect of the CDI program on the acreage of local maize (Columns 1-3), hybrid maize (Columns 4-6) and soy (Columns 7-9) in the 2014-15 season. We find that the CDI program significantly increases the hybrid maize acreage among the non-CDI club members in villages with CDI clubs, and increases the soy acreage among both non-CDI club members and CDI club members in villages with CDI clubs. The effects sizes are significant, around 0.2-0.3 acres (which is 20% in the case of maize, and 50% in the case of soy). It is interesting to note that, although not statistically significant at the 10% level, that the acreage of local maize decreases in the CDI club villages, among both CDI club and non-club members, possibly indicating a reallocation of land from local to hybrid maize and/or soy.

Table 17 presents the average treatment effect of the CDI program on the price of soy in the 2014-15 season. As CDI's market intervention did not cover the first season (2014-15) and only started in the 2015-16 season, we did not expect any effects yet for this output variable. The results confirm this hypothesis. We expect to see an effect on the prices in the endline analysis – in particular an increase in producer prices received for soybean and groundnut.

Table 18 presents the average treatment of the CDI program on the yield of hybrid maize (in kg/acre). We report no statistically significant effects on yields of hybrid maize. Recognizing a significant measurement error in this variable due to the common intercropping arrangements (we ask farmers to 'guess' the acreage of land), we do note that this non-significant result might be because the main learning event – the farmer field day – took place at the end of the season. Hence, we would expect farmers to only apply what they have learned the following season. (Note however that this appears to be at odds with the increased acreage and change in fertilizer use though). Hence, it makes sense to look at subjective yield beliefs as these will capture what farmers have learned and intend to apply the following season. The results are presented in Table 19. Columns (1) through (6) present the effects on hybrid maize yield (with columns (3)-(6) including control variables). Columns (7) through (12) present the effects soy yield (again columns (9) through (12) include control variables). We see that the CDI program increases the subjective yield beliefs of hybrid maize for the non-club members living in villages with CDI clubs by, on average, 5.3 50-kg bags (shelled), which corresponds to an effect size of 17%. The CDI program also increases the subjective yield beliefs of soy for, again, non-club members living in villages with CDI clubs, by, on average, 2.2 50-kg bags (shelled), which corresponds to an effect size of 18%. Inclusion of baseline yield beliefs (among the set of control variables) alters these results: The baseline yield expectations (despite the fact that the baseline yield for maize did not distinguish between local and hybrid maize) are highly correlated with the midline yield expectations and wipe out the treatment effects. These results a little concerning and require further investigation: as the baseline results confirmed that the randomization was successful (i.e., the balance table revealed no statistically significant differences), we suspect that the culprit might be the construction of the comparison group. We propose to test for balance at baseline among these constructed comparison groups and, if these tests reveal issues, try alternative methods to construct the comparison groups.

Overall, one notes that the treatment effects appear to be concentrated among the non-club members in the villages with clubs. Club selection might explain this phenomenon: It is possible that club members already are somewhat more progressive farmers, in the sense that they already apply the latest techniques and are generally more knowledgeable. Their higher education level points at this. As such, one might expect the largest effects among the non-CDI club members whom can receive information by passing by the demonstration plot, talking to CDI club members, and in some cases, attending the CDI field-days.[8] To check this hypothesis, we look at some additional outcomes: in particular the ones that we consider to represent quite new techniques that even the more progressive farmers might not be aware of. Table 20 presents the results on soy inoculation, a technique very much promoted by CDI. We see that the CDI program, on average, increases the adoption of this practice by 8 percentage points among CDI club members in villages with CDI clubs. This represents an effect size of close to 80%. We also (again) note a statistically significant effect on the farmers who live in treatment villages which did not form a CDI club, which is a little odd (we will return to this issue below).

In Table 21, we present the results on the midline knowledge test, disaggregating by question, this time only for the CDI club members. We see that the CDI program increases the knowledge of questions which related to chemicals – question 4: which chemical is used to control for soya rust – is answered significantly better by the CDI club members after the CDI program. Question 10 – which pesticide is used to control for cutworms – in a similar vein –is also answered more accurately by the CDI club members after the CDI program. Note however that there is a negative effect on some questions: these questions pertain to the timing of urea fertilizer and fertilizer trees (see footnote 8 as well – it might be possible that the demonstration plots did poorly in these dimensions).

In addition, it is interesting to note that we are picking up treatment effects among the farmers who live in treatment villages that do not have CDI clubs. These effects, if not an econometric artifact[9], might be due to the fact that at the time of both sensitization and lead farmer training – where the main ideas of the program were laid out – several villages who did not end of forming clubs (including some control villages) attended.

4.4. Spill-overs between the treatment and control group

As noted in the baseline report, we witnessed spillovers between the treatment and the control villages especially during the first few months of the program. The participation data at midline also suggests some spillovers even at the CDI field-day stage (which was in April). While the spillovers past December 2015 were very limited (only 10 people from control villages attended the field days), we intend to use the baseline network data and also the GPS village coordinates (see agronomic report) to study this possibility in future analysis to the extent possible.

[8] Alternatively, and more concerning, it might be possible that the CDI program has no, or even a negative effect, on the club members, as the club-members often first hand learn from their demonstration plot. As many demonstration plots had issues – see agronomic report – they might have come away with a different message than CDI intended. In future analysis, we intend to focus on the effect of demonstration-plot based learning among club members using a difference-in-difference approach and details on the demonstration plot input use and performance.

[9] The number of clusters in this sub-sample is very low, and hence a clustered regression might not be appropriate. We hence propose to repeat the analysis using robust standard errors.

5. Discussion and future plans

These first results at midline are exciting. Our analysis has made both methodological and empirical contributions to the literature and we expect the results to be relevant to policy discussions on farmer engagement and extension.

CDI's Anchor Farm Project (farmer groups, demonstration plots, training, farmer field days) is similar to a standard farmer-group based extension model implemented in developing countries by governments and NGOs. Because of the typical complexity, scale, and village-level implementation, this model for reaching and building the capacity of farmers has not received much careful attention and evaluation in the literature. Thus we are evaluating a model for farmer engagement that is common in the developing world but which has received little rigorous analysis. Our project is already providing critical insights, unpacking the effect of farmer groups from exposure to demonstration plots and participation in farmer field days on farmer learning and beliefs as we consider several important questions. First, how can we explain heterogeneity in governance and decision making across farmer groups and how is this variation related to public goods provision? Second, what is the role of farmer group functioning and endogenous farmer group characteristics related to decision-making and governance in farmer technology adoption and learning? Third, what spillovers from farmer groups and farmer field days can projects anticipate and what are the pathways through which these spillovers operate?

Relatedly, the project presents an opportunity to work out important complexities around participation in project activities. The CDI Anchor Farm Project has numerous components and numerous ways to define farmer participation. As discussed, farmers exhibited variation at midline with respect to the portfolio of activities that they took up. Some farmers participated in field days; some visited demonstration plots; some attended farmer group meetings. We are excited about the insights that the endline data will be able to provide on questions of participation. Do farmers move in and out of project activities? Do they try an activity once and then stop? Do they increase involvement over time? What farmer characteristics are associated with selecting in and out of various project activities? How does club functioning and formation evolve in time? Do all members stay for the project duration or do clubs exhibit attrition out of activities over time? Who stays and who quits? Farmer club dynamics are critical for project success and sustainability and yet we can find no published studies on this topic. These are questions that are rarely examined in the literature and we are eager to explore them using the endline data.

In addition, endline data will allow us to perform a longer term assessment of project impact on farmer adoption, beliefs, and learning. CDI is beginning a marketing component of the project this year that we are excited to evaluate in conjunction with the production and technology adoption activities. The endline sample is larger than the midline as we will return to all original 2500 farm households. This larger sample will allow us to perform heterogeneity analysis, studying the effect of the project components on female-headed households for example and on households with variable starting soil quality. It is with this data that we will be able to assess the full impact of CDI project activities and to provide insights on the broader suitability and relevance of the CDI model.