Information, credit, and inputs: the impacts and mechanisms of a program to raise smallholder productivity*

Joshua W. Deutschmann[†] Maya Duru [‡] Kim Siegal [§] Emilia Tjernström[‡]

July 22, 2018

Preliminary work; please do not cite

Abstract

Raising smallholder agricultural productivity has the potential to boost GDP and to reduce rural poverty, but the evidence on how to best achieve productivity gains remains mixed. Technologies exist that can increase smallholder yields and profits, but researchers and practitioners grapple with how to induce technology adoption in this population. This task is particularly challenging and pressing in sub-Saharan Africa, where agricultural yields lag all other regions and where poverty is often concentrated in rural areas. This paper presents randomized evidence of a rare scaled-up success story. Analyzing data from a pre-registered randomized control trial, we show that participation in One Acre Fund's (1AF) small farmer program causes statistically and economically significant increases in yields and profits. The program is designed around the notion that farmers face multiple constraints simultaneously, so that relaxing a single constraint may be insufficient. Participating farmers receive flexible financing for the purchase of high-quality seeds and fertilizer, crop insurance, and information about improved farming practices. To better understand the mechanisms through which the program works, we examine a sample of farmers who participated in the program in previous years and who use many of the recommended practices but fail to achieve yield gains similar to current participants. This suggests that relaxing information constraints alone may not be sufficient and additional results imply that the program's credit component may be crucially important. We further explore potential impact heterogeneity using simple binary interactions and new machine learning techniques allowing for a flexible functional form. Here, we observe that small farmers—who tend to be more credit-constrained—benefit more from the program. Our work has implications for poverty traps theory, as well as for the optimal design of social policy.

JEL codes: O12, O13, Q12

*Comments welcome: jdeutschmann@wisc.edu. We thank Brad Barham, Laura Schechter, UW AAE Development Lab members, and participants at IDEAS Summer School in Development for helpful comments and suggestions. All remaining errors are our own.

[†]University of Wisconsin-Madison
[‡]J-PAL
[§]One Acre Fund

1 Introduction

Raising smallholder productivity remains a key development challenge, and it has proven particularly intractable in sub-Saharan Africa. In 2014, US agricultural productivity exceeded African productivity by 473% (Magruder, 2018), and rural poverty remains much higher than urban levels (Beegle et al., 2016). The potential economic gains are significant: Gollin et al. (2018) estimate that a 10 percentage point increase in adoption of high-yielding crop varieties is associated with a 10-15 percentage point increase in GDP per capita. Furthermore, GDP growth originating in agriculture benefits the poor substantially more than growth originating in other sectors (Ligon and Sadoulet, 2008).

Agricultural policy experiments such as extension programs, fertilizer subsidies, and information interventions—tested by governments, NGOs, and researchers alike—have yielded substantial evidence on what fails and a limited catalogue of successes. Is this limited progress a sign that African agriculture is a lost cause or do existing programs lack something crucial? This paper uses a pre-registered, randomized control trial to examine the One Acre Fund (1AF) small farmer program, which succeeds in boosting smallholder productivity. By providing farmers with training, in-kind credit, insurance, and access to high-quality inputs, the program increases per-acre farmer yields by 16% and profits by 21% in our preferred specification.

The 1AF program's focus on credit, insurance, inputs, and information is no accident. The economics literature has accumulated substantial evidence that failures in any these domains hinder farmers' ability or willingness to adopt improved agricultural technologies. Feder et al. (1985) emphasize three of these constraints in their review of the technology adoption literature at the time (credit, risk, and information). Magruder (2018) focuses on the same three constraints as key barriers to adoption in a more recent review of the experimental evidence on technology adoption. The focus on providing high-quality inputs also has empirical support: a growing literature establishes that fertilizer available on local markets in the region often falls short of quality standards (Bold et al., 2017; Fairbairn et al., 2017; Tjernström et al., 2018).

Input loan interventions on their own have seen relatively low levels of takeup, and have had limited impacts on crop choices and productivity (Giné and Yang, 2009). Ashraf et al. (2009) examine a program that combines credit with increased access to export markets and find that bundling the two increased program takeup but failed to increase farmer incomes relative to farmers who were only offered the export program. Despite facing clear liquidity constraints, few farmers took up a voucher program for fertilizer and improved seeds in Mozambique (Carter et al., 2013). However, the farmers who did experiment with fertilizer in response to the vouchers continued to use the technology even after subsidies expired, suggesting a role for learning about optimal fertilizer use (Carter et al., 2014).

Theory clearly suggests that uninsured risk reduces input use, which in turn implies that crop insurance should increase input use. Empirical tests of this hypothesis show that when farmers receive insurance, they indeed increase their input use and shift towards higher-yielding but riskier crop varieties (Karlan et al., 2011; Mobarak and Rosenzweig, 2012; Elabed and Carter, 2014). Some studies suggest that insurance may be a more important driver of agricultural investment than credit (Karlan et al., 2014). Despite the promise that insurance holds, a large number of index insurance trials in a wide variety of contexts have been faced with very low take-up. We are unable to directly investigate the importance of the insurance component of 1AF's program, but will keep in mind the fact that farmers have access to crop and funeral insurance in our discussion of mechanisms below.

In terms of information, farmers face a stochastic environment that makes it difficult to infer what optimal practices and technologies are. Even beyond the information needed to make adoption decisions about new technologies, farmers may persist in using sub-optimal agricultural practices because the noisiness of agricultural yields hampers their ability to learn. Providing detailed information about specific ways to improve yields has been found to be effective at shifting farmer practices (Hanna et al., 2014). However, even if farmers do adopt improved practices, studies have found that they sometimes make small mistakes that render adoption unprofitable (Duflo et al., 2008). Furthermore, farmers tend to learn and extrapolate from both their own experiences (Maertens et al., 2017) and those of their neighbors (Conley and Udry, 2010). This can help diffuse information through networks, but it may also perpetuate flawed information, especially when outcomes are highly stochastic and dependent on unobserved factors.

Beyond knowing that the program works, we would also like to disentangle the mechanisms through which it works and to better understand the extent to which its bundled nature contributes to its success. A growing body of evidence shows that bundled interventions can be effective at moving people out of poverty (Bandiera et al., 2017; Banerjee et al., 2015). If we believe that farmers face multiple simultaneous constraints, offering an intervention that only relaxes credit or only provides information may not be sufficient to raise yields and profits in a significant or economically meaningful way.

To make progress towards the goal of understanding mechanisms, we therefore perform a number of additional analyses: first, we examine a sample of farmers with prior (non-randomized) access to the program. These farmers continue to use of many of the improved agricultural practices that the program teaches, but these practices do not translate into statistically significant yield gains. We interpret this as suggestive evidence that relaxing information constraints may not be sufficient on its own. Second, heterogeneity analysis using both binary interactions and new machine learning techniques reveals that small farmers—who tend to be more credit-constrained—benefit more from the program. Third, a comparison of farmers' fertilizer use on program plots and non-enrolled plots corroborates the notion that the program's credit component may be central to its success. While the RCT that we study did not randomly vary the components (as in Banerjee et al. (2015)), we believe this analysis sheds additional light on the program mechanisms.

The rest of the paper is organized as follows. Section 2 describes the context of our study, the experimental design, and the different samples that we employ in our analysis. This section also familiarizes the reader with our sample by presenting summary statistics and balance checks. Section 3 describes our empirical analysis—both the pre-registered RCT analysis and the more exploratory investigations of mechanisms. Section 4 presents our primary empirical analysis of the effects of program participation on maize yields and profits, and discusses the improved practices that treated farmers adopt. Section 5 then investigates why pre-exposed farmers do not seem to be able to obtain the higher yields that we see among treated farmers, despite using many of the improved practices. We build on this examination to analyze whether information and credit seem to be binding constraints. Section 6 conducts a host of robustness checks of our main results, and Section 7 concludes.

2 Context and data

This study focuses on Kenya, where forty percent of the population lives in poverty, and a large share of the poor engage in smallholder farming as their primary occupation. The agriculture sector contributes 51 percent to the country's GDP (25% indirectly) and is dominated by small-scale producers (Government of Kenya, 2010). Despite the importance of the agricultural sector to the economy, most smallholder farmers are not running successful micro-enterprises. Households in Kenya typically derive their income from the production of a variety of crops, often combined with a range of off-farm activities (Sheahan et al., 2013). Most smallholder farmers produce less food than they need to feed their families, and are net maize buyers (Kirimi et al., 2011).

We further focus on maize—the main staple crop in Kenya—as the crop is crucial both to the economy and from a food security perspective. Seventy percent of Kenya's maize is produced by smallholders who farm between 0.2 and 3 hectares (Government of Kenya, 2010, pg. 11-12). While Kenyan farmers use improved improved maize varieties and inorganic fertilizer at higher rates than other countries in the region, maize yields remain low. Increasing yields and profits is crucial if we want the agricultural sector to act not simply as a means of subsistence, but as a pathway out of poverty.

1AF has operated in western Kenya for more than ten years, and served over 200,000 farmers in 2016. Their core "market bundle" addresses several of the key constraints discussed in Section 1: participating farmers receive group-liability loans for improved seeds and high-quality fertilizer, weekly training on modern agricultural techniques, crop and funeral index-based insurance, and market facilitation support to help them sell their products for higher prices.

Farmers choose the amount of land to enroll, and 1AF provides the agricultural inputs as a function of the amount of land enrolled. The group liability loans are given to self-selected farmer groups. The loan terms are quite flexible, allowing farmers to repay in any amount at any time during the growing season (but they must complete repayment in full by the end of the harvest). Historically, repayment rates have been over ninety-seven percent. Field officers conduct the weekly trainings with farmer groups; a field officer interacts on average with around 200 farmers.

2.1 Experimental design

The experiment was carried out in a cluster-randomized design in the Teso region of Kenya. As described above, farmers self-select into farmer groups of around 10-12 farmers. A cluster is then made up of 2-4 of these joint-liability farmer groups. Participants were recruited following standard 1AF protocol. Typically, once a farmer indicates interest in signing up and satisfies the basic criteria, he/she pays a small program participation deposit of approximately \$5 USD.[1] For the purposes of the study, farmers also had to consent to take part. Shortly after contract signing, 1AF informed farmers about the study, that their participation would be voluntary, and further provided them with informed consent documents. 1AF informed farmers that half of them would be randomly assigned to treatment, while the other half would receive an alternative compensation package (household goods and a discount for 1AF participation the following season amounting to roughly 20% of the typical program cost). Randomization was conducted by public lottery, and all farmers enrolled in the study also received 10,000 Ksh in funeral insurance coverage, regardless of treatment status.[2]

The study is complicated by the fact that it took place in a region where 1AF has already operated for several years. The specific villages selected for study inclusion had never been offered the 1AF program, but neighboring villages had previously been offered the program. Thus, a substantial portion of farmers who expressed interest in participating had previously participated in the 1AF program (by "commuting" to the neighboring villages to participate). 1AF typically targets smallholder farmers living in areas with 100-500 people per km^2, who rely mostly on agricultural income, but have limited access to farm inputs. Therefore, when 1AF decided to conduct an RCT to evaluate their work in Kenya, choosing completely new regions would have forced them to expand into regions that are quite unrepresentative of the typical program. For example, they would have had to enroll farmers in pre-dominantly tea-growing areas, or located in substantially different agro-ecological zones. Teso district in particular was chosen as a suitable region, being representative of 1AF program areas, while still having many farmers in the district who had not been previously exposed to 1AF.

Because of the "contamination" of some of the sample, we define two different samples for the analysis: the "primary sample," which consists of treated and control farmers who had never previously participated, and the "full sample," which includes the pre-exposed farmers. In some of our analyses, we also separately consider only the pre-exposed sample. Tables and graphs clearly note which sample we are using. It is worth keeping in mind that both of the samples present some

[1] To join 1AF, a farmer must satisfy the following restrictions: have a phone number and national identification; agree to repay their loan; and complete pre-payment of 500 Kenyan Shillings (Ksh, equivalent to approximately \$5 USD).

[2] For the study, farmers also had to fulfill two additional requirements: they had to cultivate maize, and they had to be able to cultivate at least a quarter of an acre of maize.

challenges and some advantages. On the one hand, if the pre-exposed control farmers continue to benefit from their prior program involvement even after quitting the program (by, for example, continuing to use the new practices that they've learned), then treatment effects estimated with the full sample would likely result in a downward-biased impact estimate. On the other hand, we can examine the pre-exposed control farmers to see whether they achieve similar yields to the newly-enrolled treated, and we can compare pre-exposed farmers in the treatment group to see whether effects accumulate with multiple years of treatment.

Furthermore, the farmers who had never been exposed may be different than those who self-selected into the program in earlier years. If selection into the program is positively correlated with potential returns to the program (i.e., higher-return farmers opt in earlier), we would expect the impacts on "hold-out" farmers to be a lower bound of the true impacts; if timing of enrolment is uncorrelated with returns to program participation, then we will obtain unbiased estimates of the program.

2.2 Data collection

The data collection for this experiment was directly managed by 1AF. To increase confidence in the process, they contracted with the International Initiative for Impact Evaluation (3ie) to help design and review all parts of the trial—including the experimental design, the field protocols, sampling and randomization, as well as the data collection instruments (Dubey and Yegbemey, 2017). 3ie concluded that the randomization was conducted successfully, and noted that 1AF staff showed high level of professionalism in conducting the randomization.[3]

Readers may also worry about the potential for enumerator effects since 1AF collected the data from farmers. This arrangement was agreed upon for practical purposes as 1AF has dedicated and trained staff on the ground to collect the large amounts of household and physical harvest data that were required for this study. Several factors should help inspire confidence in the data collection: first, 1AF used standard best-practice protocols to ensure data quality, such as back-checks (resurveying 10% of respondents and comparing with original results) and in-field supervision. Second, 1AF contracted an independent survey firm, Intermedia Development Consultants (iDC), to carry out

[3] A letter from 3ie attesting to their review and approval of the pre-analysis plan can be found in Appendix **X**.

a three-step audit of the data collection[4]. The audit team participated in group meetings held by the 1AF monitoring team, spent 26 days in the field, during which they observed the work of enumerators on 76 occasions and enumerators' interactions with farmers on 246 occasions. They observed the work of supervisors on 11 occasions and that of supervisors as back-checkers on five occasions.(Intermedia Development Consultants, 2017) They further carried out parallel data collection efforts to the 1AF data collection, comparing the results and finding minimal discrepancies.

Two of the authors on this paper were brought in to review the pre-analysis plan (PAP) prior to follow-up data collection and to conduct the final analysis according to the PAP. The team carried out all the variable construction and analysis from scratch, i.e. without referring to 1AF's existing analytical code.[5] This paper extends the basic impact evaluation analysis from the report in order to investigate impact mechanisms; Table **Y** in Appendix **Z** clearly delineates which analyses were pre-registered and which fall outside of the purview of the PAP.

Baseline data collection occurred in November and December of 2016—after program enrollment but before treatment assignment. The public lottery, which assigned clusters of farmer groups to treatment, took place in January 2017. Enumerator teams rolled out input use surveys after the planting of the main season in 2017, from April through June. These data provide detailed information on the extent of farmer compliance with treatment, such as whether farmers spaced their plants correctly, and applied the correct fertilizer dosage. We use these data as indicators of the extent to which farmers are actually learning and changing their behaviors as a result of the 1AF training.

The main outcome variables come from a harvest data collection effort that 1AF carried out during the main maize and beans harvest period (May to October, 2017). The procedure for collecting yield data is the following: for treatment farmers, enumerators randomly selected one maize plot that was enrolled in the program ("enrolled plot") and one maize plot that was not enrolled ("non-enrolled plot"). For control farmers, enumerators randomly selected a random maize plot. Enumerators then collected wet and dry harvest weights from two randomly placed 8 x

[4]The iDC audit report is available from the authors upon request; the overall conclusion of the audit stated that "With respect to the Teso trial, the strategy and planning are appropriate to the situation, and they have attempted to make every effort to obtain accurate, reliable and valid results. ... Overall, the data collection is well planned and executed. The possibilities for improving performance are quite limited. The survey team members are well selected and committed to the process. The recording and transmission of data is well done, with minimal errors."

[5]The final report can be found here: Deutschmann and Tjernström (2018) [link to report]

10 boxes for each selected plot. These estimates are applied to all cultivated land to determine per-farmer yields (for treatment farmers, we compute a weighted average of 1AF and non-1AF cultivated land size). We expect this method to produce substantially better yield estimates than the self-reports commonly reported by economics studies. Furthermore, we measure land size based on GPS readings of each field (we can also compare this to self-reported land size; more on this in Section 6).

2.3 Summary statistics

Table 1 shows baseline means of key control variables in the full sample, as well as balance tests across treatment and control samples. The majority of the sample is married, and less than half of male household heads have completed secondary school. The majority of households (just under 80%) earned more than 50% of their income from farming in the last year, corroborating that farming is a key activity in this population. The average farmer plants one acre of maize and gets roughly half a ton (500 kgs) of maize per acre.

Many households use some type of agricultural technology at baseline (roughly 77%), but this basic variable does not capture anything about the intensity of agricultural technology use. Figures 1 and 2 show the baseline fertilizer intensity of two key inorganic fertilizers, DAP (diammonium phosphate) and CAN (calcium ammonium nitrate). These two graphs clearly show that while many farmers used some fertilizer at baseline (see the lines representing "new participants" in both treatment and control), they applied them at very low rates. We can also see that pre-exposed farmers tend to use more fertilizer than new participants, suggesting that they learned something from 1AF participation and/or that earlier enrollees are better farmers to start with.

2.3.1 Full sample, baseline balance

Only two variables are statistically different across the two groups: previous knowledge of 1AF practices, and self-reported per-acre yields. Although the differences in the other covariates are not themselves statistically distinguishable from zero, the F-test of joint orthogonality is significant at the 5%. We include all of the variables in this table as controls for all regressions (as per the PAP), unless otherwise noted.

Variable	(1) Control N	(1) Control Mean/SE	(2) Treatment N	(2) Treatment Mean/SE	T-test Difference (1)-(2)
Married	1190	0.879 (0.009)	1237	0.873 (0.009)	0.006
Father, 2ary school (0/1)	1199	0.395 (0.014)	1230	0.426 (0.014)	-0.031
Farm labor >50% income	1190	0.772 (0.012)	1237	0.758 (0.012)	0.014
Used ag tech 2016	1199	0.762 (0.012)	1257	0.778 (0.012)	-0.016
Prev. 1AF knowledge	1190	0.456 (0.014)	1237	0.517 (0.014)	-0.060***
Intercropped 2016	1190	0.477 (0.014)	1237	0.468 (0.014)	0.009
Credit access 2016	1190	0.709 (0.013)	1237	0.728 (0.013)	-0.019
Maize acres, 2016	1190	1.007 (0.024)	1237	1.006 (0.022)	0.001
Household size	1190	6.648 (0.073)	1236	6.780 (0.075)	-0.132
Maize yield/acre, 2016	1180	530.242 (12.463)	1213	580.555 (13.403)	-50.312***
F-test of joint significance (F-stat)					1.985**
F-test, number of observations					2345

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. Fixed effects using variable siteid are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 1: Baseline balance of treatment and control samples

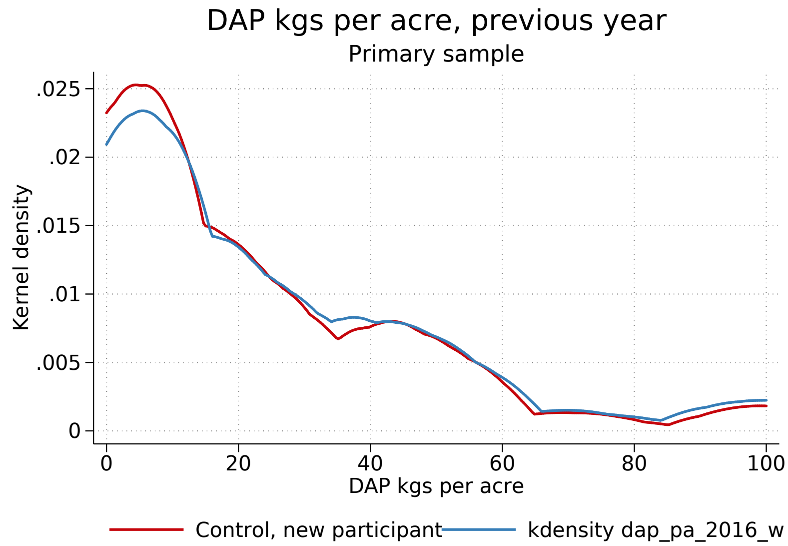

Figure 1: Fertilizer use at baseline (DAP)

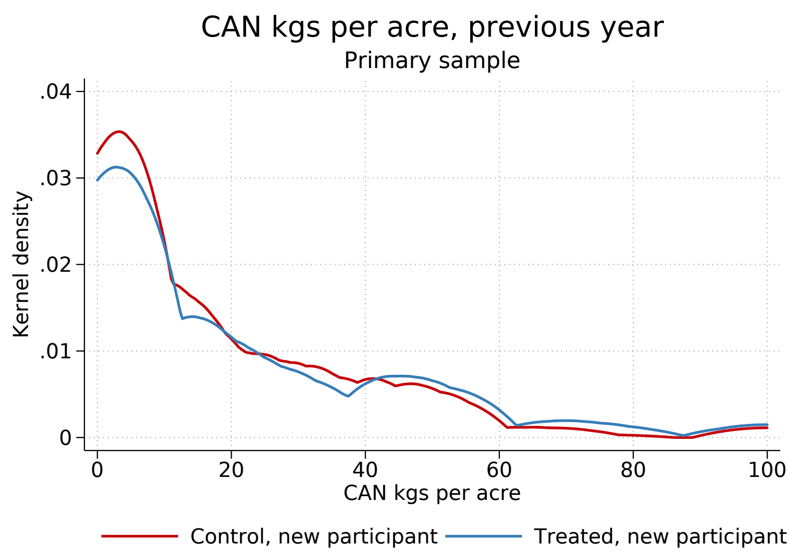

Figure 2: Fertilizer use at baseline (CAN)

2.3.2 Pre-exposed vs. new participants

Table 2 compares the farmers across the pre-exposed and new participant samples. As Figures 1 and 2 already hinted at, the two samples differ substantially. Pre-exposed farmers are more likely to have an educated father in the household, more likely to earn at least 50% of household income from farm labor, more likely to report using improved seeds or fertilizer at baseline, more likely to report baseline knowledge of 1AF practices, less likely at baseline to intercrop beans with maize, more likely to report having access to credit, reported using more land for maize at baseline, have bigger households, and reported higher maize yields per acre at baseline.

Several of these differences (education, land endowments, yields, etc.) suggest that farmers who enrolled earlier may be better farmers—although several of these variables were likely directly affected by program participation. Because of this, in most cases we prefer to present results using the primary sample, despite the loss in power afforded by the smaller sample size. However, our results are robust to using either sample definition.

3 Empirical strategy

This section discusses our analytical approach, both for the primary analysis (as pre-specified in the PAP) and for the additional analyses that try to dig into impact heterogeneity and mechanisms.

3.1 Pre-registered main analysis

The main analysis focuses on three outcomes of interest: per-acre maize yields, per-farmer maize yields, and profit (in USD). Maize yields use dry weight measurements from randomly-selected areas on farmers' plots. For treated farmers, the yield outcomes are a weighted averages of yields on enrolled and non-enrolled land. Profits account for program costs, input costs, labor costs, and average local market prices over the off-season. Revenues for profit calculations account for both maize and any inter-cropped bean yields, since around half of the sample farmers intercrop (it is difficult to separate labor or fertilizer costs across crops within the same plot). See Appendix A for a more detailed explanation of how the outcome variables are constructed.

	(1)		(2)		T-test
	Primary Sample		Pre-Exposed Sample		Difference
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)
Married	874	0.886	1553	0.871	0.015
		(0.011)		(0.009)	
Father, 2ary school (0/1)	896	0.363	1533	0.439	-0.076***
		(0.016)		(0.013)	
Farm labor >50% income	874	0.757	1553	0.769	-0.012**
		(0.015)		(0.011)	
Used ag tech 2016	920	0.591	1539	0.876	-0.285***
		(0.016)		(0.008)	
Prev. 1AF knowledge	874	0.097	1553	0.706	-0.609***
		(0.010)		(0.012)	
Intercropped 2016	874	0.509	1553	0.452	0.057***
		(0.017)		(0.013)	
Credit access 2016	874	0.706	1553	0.726	-0.020***
		(0.015)		(0.011)	
Maize acres, 2016	874	0.955	1553	1.035	-0.081***
		(0.027)		(0.021)	
Household size	874	6.597	1552	6.782	-0.184***
		(0.087)		(0.066)	
Maize yield/acre, 2016	860	433.120	1533	624.538	-191.418***
		(13.794)		(11.690)	
F-test of joint significance (F-stat)					117.626***
F-test, number of observations					2345

Notes: The value displayed for t-tests are the differences in the means across the groups. The value displayed for F-tests are the F-statistics. Fixed effects using variable siteid are included in all estimation regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table 2: Balance across primary and pre-exposed samples

For each outcome, we estimate the following regression:

$$Y_{is} = \alpha + \beta T_{is} + \delta X_{is} + \gamma_s + \epsilon_{is} \tag{1}$$

where T_{is} is the treatment dummy for individual i in field officer area s, X_{is} includes the list of pre-specified controls, γ_s is a field office area fixed effect, and ϵ_{is} is clustered at the cluster (treatment assignment) level.

To ensure correct inference given our multiple hypotheses, we follow Jones et al. (2018) and implement Westfall and Young (1993) free step-down re-sampling to control for family-wise error rate. We treat each regression table as a family of hypotheses. We present adjusted p-values based on this method at the bottom of each table in Section 4. The method does not change the significance level of any of the treatment coefficients.

3.2 Heterogeneity analysis

For the first part of the heterogeneity analysis, we estimate a regression in which we interact the treatment dummy with binary indicators of various baseline characteristics, as in the following:

$$Y_{is} = \alpha + \beta_1 T_{is} + \beta_2 K_{is} + \beta_3 (T_{is} \times K_{is}) + \delta X_{is} + \gamma_s + \epsilon_{is} \tag{2}$$

where K_{is} is the key variable of interest for heterogeneity analysis (for individual i in field officer area s). Most of our analysis will show the point estimate and 95-percent confidence interval of the average marginal effect of treatment for farmers with farmers with $K_{is} = 1$ and $K_{is} = 0$.

For a second set of heterogeneity results, we employ a Generalized Random Forest (GRF) methodology. GRF is a non-parametric machine learning algorithm that can help identify which (if any) baseline characteristics are associated with greater program impacts. Broadly speaking, a random forest is a collection of classification trees that are built using data and a statistical learning algorithm. Each classification tree is created by randomly drawing a subset of observations, called the training sample. The training sample is then drawn without replacement from the full sample, and forms the basis for a tree. The training sample is recursively split into subgroups along values of some covariates. We can then predict an individuals outcome, y_i, using the mean y of observations that share similar covariates. In turn, the average of these predictions across many trees is an estimate of the conditional mean of y. The algorithm determines which splits that are implemented based on those that minimize some in-sample goodness-of-fit criterion.

The Generalized Random Forest (GRF) method developed by Athey, Tibshirani and Wager (2018) improves upon random forest methods by proving that forest estimators are consistent and have an asymptotic normal distribution with a variance that can be estimated. In the context of heterogeneous treatment effects, the GRF algorithm can be used to build a causal random forest (CRF) to estimate a conditional average treatment effect. We report graphical results representing the predictions of treatment effects as we vary certain baseline variables, holding all other covariates at the sample median.

Future work in progress includes analysis following the recently developed machine learning methods in Chernozhukov et al. (2018).

4 Program impact

This section presents our main results on program impact, in which we evaluate the impact of the 1AF market bundle on yields and profits. Table 3 shows result using our preferred specification and the smaller (not pre-exposed) sample. Column (1) shows the impact of the treatment on per-acre maize yields, column (2) shows effects on per-farmer yields, and column (3) shows impacts on profits. These correspond to a 16% increase in per-acre yields, a 32% increase in per-farmer yields, and a 20% increase in profits.

The magnitude of the profit results is particularly remarkable since even programs that affect yields often fail to increase profits. Further, treatment farmers spent a mean of $112 USD (median $98) on inputs, labor, and program participation costs, compared to a mean of $76 (median $63) for control farmers. Using this primary sample estimate of the treatment effect on profits, this suggests a 65% return on average for farmers who participated. Similarly, if we compare the estimated coefficient with self-reported income at baseline (mean $751, median $481), this increase in profit is far from trivial.

While we report our preferred specification here (including covariates, site fixed effects, and using the primary sample), we additionally conduct a number robustness checks in Section 6, hopefully convincing even skeptical readers that our results do not hinge upon these model assumptions. We demonstrate that these results are robust to different sample definitions and to the exclusion of the covariates and of site fixed effects. Also reported in Section 6) are results using randomization inference, inspired by Young (2017).

We also examine whether the treatment affected farmers' practices. Table 4 shows results examining the effect of 1AF participation on various improved farming practices. Column (1) reports the effect on the number of acres planted in maize. The results suggest that the program is not substantially affecting how much land farmers are devoting to maize, confirming that the per-farmer yield impacts from Table 3 are not coming simply from an increase in planting area. Columns (2) and (3) look at the treatment impact on farmer likelihood of using a commercial seed and the intensity of improved seed use. The program affects neither of these practices, despite commercial seed being one of the main components of the 1AF bundle. This could be due to high baseline use of commercial seed among both treatment and control farmers. Column (4) shows treatment impacts on the likelihood of intercropping with beans, a practice that 1AF recommends. The lack of impact here could be due to the large share of farmers who were already adopting that practice at baseline.

Columns (5) - (8) examine fertilizer use. DAP (diammonium phosphate) is a fertilizer that should be applied at the time of planting. It is the fertilizer type that farmers are most likely to be familiar with. However, anecdotally some farmers incorrectly apply it also later in the season, which tends to be ineffective. CAN (calcium ammonium nitrate) on the other hand is less commonly used in this area. It should not be applied at planting, but should instead be applied when the maize plant is about knee-high. Here, we do not examine application rates, but rather the timing of fertilizer application. In columns (5) and (6), we can see that the treatment increases the likelihood that a farmer applies DAP fertilizer at planting, but has no effect on the at-planting application of CAN. This is encouraging, as it seems farmers are applying the fertilizer at the recommended time. The post-planting behavior in columns (7) and (8) also suggest that treated farmers are more likely to apply CAN and to do so in the recommended manner (post-planting). We will examine fertilizer rates when we turn to mechanisms in Section 5.

We can also examine several other recommended practices, including whether the farmer prepared her field with a plow, and whether the spacing of plants and rows within the field are according

Table 3: Primary outcomes, primary sample

	(1) Maize PA	(2) Maize Per Farmer	(3) Profit
1AF participant	181.673***	345.074***	72.677**
	(35.821)	(84.088)	(30.418)
Married	106.281*	-41.144	-23.607
	(59.095)	(152.252)	(56.302)
Father, 2ary school (0/1)	-4.903	249.461***	88.690**
	(45.756)	(94.245)	(35.212)
Farm labor >50% income	8.930	-10.724	-1.115
	(49.994)	(84.821)	(31.249)
Used ag tech 2016	-69.550	-100.434	-44.421
	(43.968)	(86.068)	(31.066)
Prev. 1AF knowledge	14.105	114.447	33.869
	(62.882)	(166.821)	(62.305)
Intercropped 2016	-5.946	-138.885	-48.496
	(45.297)	(94.922)	(34.708)
Credit access 2016	-79.429**	-65.529	-26.848
	(38.925)	(84.676)	(30.536)
Maize acres, 2016	57.257**	413.300***	132.809***
	(27.106)	(79.716)	(28.756)
Household size	3.421	35.935**	13.455**
	(7.267)	(17.200)	(6.036)
Maize yield/acre, 2016	0.177***	0.461***	0.155***
	(0.051)	(0.108)	(0.040)
FAW Incidence	-6.352*	-1.091	0.470
	(3.708)	(8.499)	(3.144)
Observations	701	637	637
R^2	0.136	0.282	0.241
MHT p-value for 1AF	0.000	0.000	0.019
Control Mean Dep. Var	1124.733	1081.431	353.263

Standard errors in parentheses

Note: standard errors clustered at cluster level. All regressions include field office area fixed effects. MHT p-value uses Jones et al. (2018) implementation of Westfall and Young (1993) correction for multiple hypothesis testing.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 4: Study year (2017) practices, primary sample

	(1) Maize Acres	(2) Use Comm. Seed	(3) Kg Comm. Seed	(4) Int. Beans	(5) CAN AP	(6) DAP AP	(7) CAN PP	(8) DAP PP	(9) Used Plow	(10) Row Spacing	(11) Plant Spacing
1AF participant=1	0.041	-0.003	0.028	0.018	-0.005	0.539***	0.280***	-0.146***	0.051	0.312***	0.262***
	(0.053)	(0.026)	(0.321)	(0.029)	(0.010)	(0.032)	(0.028)	(0.023)	(0.034)	(0.040)	(0.031)
Married	-0.063	0.044	0.001	0.030	0.001	-0.010	0.023	-0.002	-0.002	0.080	0.106***
	(0.078)	(0.049)	(0.366)	(0.046)	(0.016)	(0.038)	(0.049)	(0.038)	(0.042)	(0.049)	(0.038)
Household size	0.048***	0.007	0.019	-0.000	-0.002	0.007	-0.005	-0.003	0.001	-0.012*	0.002
	(0.010)	(0.005)	(0.085)	(0.006)	(0.002)	(0.006)	(0.006)	(0.005)	(0.006)	(0.007)	(0.006)
Father, 2ary school	0.232***	0.062*	0.161	-0.011	0.015	-0.004	0.079***	-0.047	0.055**	0.050	-0.058**
(0/1)	(0.046)	(0.033)	(0.409)	(0.026)	(0.010)	(0.023)	(0.030)	(0.031)	(0.026)	(0.037)	(0.027)
Farm labor >50%	0.067	0.065*	0.635*	0.046	-0.009	-0.010	0.048	0.035	-0.006	-0.045	-0.017
income	(0.053)	(0.038)	(0.354)	(0.034)	(0.013)	(0.031)	(0.031)	(0.027)	(0.040)	(0.049)	(0.031)
Maize acres, 2017		0.037*	2.465***	0.035	0.001	-0.028	0.000	0.064***	0.081***	0.014	-0.021
		(0.021)	(0.379)	(0.023)	(0.007)	(0.021)	(0.021)	(0.021)	(0.025)	(0.030)	(0.018)
Observations	841	841	835	835	827	827	833	833	827	841	841
R^2	0.157	0.044	0.159	0.320	0.023	0.394	0.167	0.219	0.268	0.128	0.174
Ctrl Mean Dep. Var	0.921	0.761	5.285	0.402	0.020	0.409	0.623	0.244	0.738	0.349	0.091

Standard errors in parentheses

Note: standard errors clustered at cluster level. All results are from OLS regressions which include field office area fixed effects. AP stands for at planting, and PP stands for post planting. 1AF recommends using DAP at planting and CAN post planting. Row spacing and plant spacing are 1 if average spacing within 5 cm of recommended.

* $p < .1$, ** $p < .05$, *** $p < .01$

to the recommended agronomic practices. Here, we define farmers as doing it correctly if their average row- and plant-spacing practices are within 5cm of the recommended method. We can see that the treatment substantially increases the likelihood of both correct row spacing and plant spacing. Since plant density tends to deviate from agronomic best-practice in most of sub-Saharan African smallholder agriculture, this improvement could have important implications for yields. Plant spacing is also a practice that should be less costly than some of the other recommended practices like applying purchased inputs like fertilizer and commercial seeds.

We additionally look at the question of improved practices as an index. We consider four key recommendations: use DAP at planting, use CAN after planting, correct row-spacing (within 5 cm of the recommended 75 cm apart), and correct plant-spacing (each plant within a row should be approximately 25 cm from the next). Here we can look at difference between treatment and control groups in terms of the number of practices that they adopt. Figure 3 shows, within the primary sample, that treated farmers indeed adopt more of the recommended practices than do control farmers, although control farmers did already tend to adopt some of the same practices.

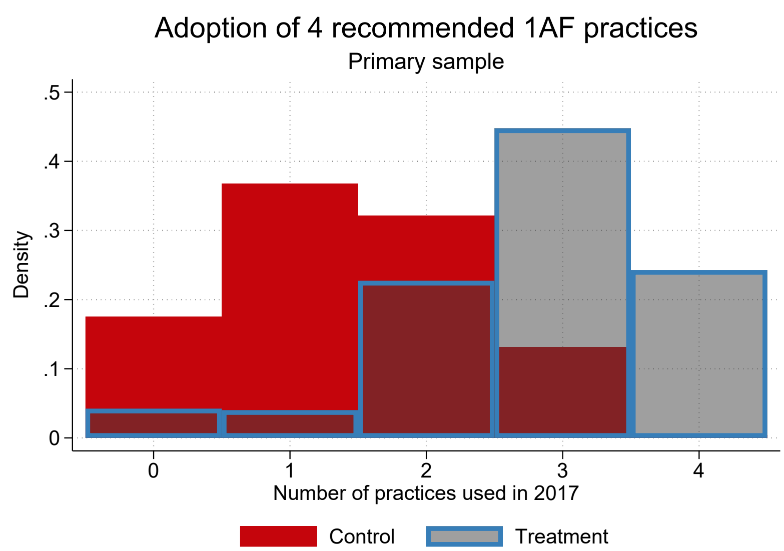

Figure 3: Improved practice adoption by treatment status

We have thus established that the treatment affects the agricultural practices of treated farmers, and that the program significantly increases yields and profits. From a policy and implementation perspective, it seems logical to wonder to what extent the training and subsequent improvements in improved practices drives yield and profit impacts. Since fertilizer and seed are costly and risky investments, adoption of low-cost improved practices may be easier to diffuse than other technologies—if indeed they drive the results.

5 Unpacking program mechanisms

We are interested in understanding how the various program components drive the results presented above. The 1AF program includes four different components: input credit, insurance, access to high-quality inputs, and training in improved agricultural practices. While none of the components were individually randomized, we will attempt to investigate the relative importance of these factors.

5.1 Information

A key component of the 1AF bundle is the training that farmers receive on optimal planting practices. As we saw in Section 4, the program does increase farmers' adherence with recommended agronomic practices, suggesting that there may have been some information gaps at the baseline. If farmers continue to apply these improved practices after they are not longer participants in the program, it suggests that they learned from the program that these practices are beneficial.

Table 5 shows suggestive evidence that pre-exposed farmers tend to use improved practices. The table only includes control farmers, and our coefficient of interest is that on the dummy for pre-exposure. Of course, since pre-exposure to 1AF's programming is non-random, it is entirely possible that farmers who selected into the program earlier were different than those who delayed enrollment. If early-adopters were better farmers, they may have been using some of the improved practices even in the absence of the program. In particular, the coefficient on past participant in column (2), which shows that pre-exposed farmers are more likely to plant a commercial seed than are non-pre-exposed farmers suggests that selection is likely a part of the story here.

With that in mind, we can note that the pre-exposed farmers plant more maize (column 1), are more likely to use DAP at planting (column 6), and correctly less likely to use DAP post-planting (column 7). They are also more prone to using a plow and to correctly space the rows of their maize (columns 9 and 10). While we cannot guarantee that the self-selected farmers did not already use these practices before enrolling, the fact that the adoption lines up quite well with the (randomized)

Table 5: Study year (2017) practices, control sample

	(1) Maize Acres	(2) Use Comm. Seed	(3) Kg Comm. Seed	(4) Int. Beans	(5) CAN AP	(6) DAP AP	(7) CAN PP	(8) DAP PP	(9) Used Plow	(10) Row Spacing	(11) Plant Spacing
Past 1AF participant	0.072*	0.112***	0.171	-0.028	-0.010	0.071**	0.056	-0.052**	0.072**	0.079**	0.016
	(0.040)	(0.022)	(0.433)	(0.024)	(0.009)	(0.035)	(0.034)	(0.026)	(0.030)	(0.035)	(0.022)
Married	0.039	0.067**	0.440	0.012	-0.018	-0.048	0.026	0.010	-0.040	0.046	0.059*
	(0.081)	(0.033)	(0.380)	(0.040)	(0.014)	(0.041)	(0.043)	(0.038)	(0.031)	(0.039)	(0.031)
Household size	0.046***	0.003	0.146*	0.004	-0.001	0.005	-0.005	-0.006	0.010**	-0.008	0.003
	(0.009)	(0.004)	(0.076)	(0.005)	(0.001)	(0.006)	(0.006)	(0.005)	(0.004)	(0.007)	(0.004)
Father, 2ary school	0.253***	0.007	0.073	-0.010	0.010	0.028	0.047*	-0.031	0.034*	0.038	0.015
(0/1)	(0.039)	(0.022)	(0.374)	(0.025)	(0.007)	(0.029)	(0.026)	(0.024)	(0.019)	(0.039)	(0.020)
Farm labor >50%	0.103**	0.068**	0.743*	0.126***	-0.006	0.026	0.032	-0.004	-0.031	-0.024	0.022
income	(0.047)	(0.027)	(0.414)	(0.028)	(0.008)	(0.036)	(0.031)	(0.026)	(0.031)	(0.036)	(0.020)
Maize acres, 2017		0.026	2.966***	0.058***	-0.004	0.019	0.033	0.038**	0.082***	0.008	-0.009
		(0.017)	(0.378)	(0.016)	(0.004)	(0.024)	(0.022)	(0.017)	(0.019)	(0.026)	(0.010)
Observations	1151	1151	1143	1143	1117	1117	1145	1145	1117	1151	1151
R^2	0.126	0.066	0.190	0.247	0.021	0.130	0.097	0.170	0.270	0.031	0.081
Ctrl Mean Dep. Var	0.955	0.856	5.412	0.471	0.014	0.493	0.689	0.240	0.765	0.392	0.119

Standard errors in parentheses

Note: standard errors clustered at cluster level. All results are from OLS regressions which include field office area fixed effects. AP stands for at planting, and PP stands for post planting. 1AF recommends using DAP at planting and CAN post planting. Row spacing and plant spacing are 1 if average spacing within 5 cm of recommended.

* $p < .1$, ** $p < .05$, *** $p < .01$

treatment effect on practices suggests that at least some of these practices are likely persistent effects of the treatment. So with all these improved practices, do pre-exposed farmers also reap the higher yields associated with 1AF participation?

Table 6 investigates this by regressing our main outcome variables on a dummy for past 1AF participation, within the control sample. As we can see, even though past participants employ several of the recommended practices, this does not appear to translate into yield or profit impacts. The point estimates are positive, but not statistically significant. To prod this a little further, we also examine whether treatment impacts vary by baseline improved practice adoption.

Table 6: Primary outcomes, control sample

	(1) Maize PA	(2) Maize Per Farmer	(3) Profit
Past 1AF participant	13.798	107.610	33.587
	(35.591)	(66.350)	(24.778)
Married	79.243	147.770	48.012
	(47.858)	(100.615)	(36.859)
Household size	1.731	50.939***	17.028***
	(6.169)	(14.495)	(5.207)
Father, 2ary school (0/1)	41.460	334.502***	115.776***
	(40.433)	(83.857)	(30.261)
Farm labor >50% income	62.216**	212.421***	80.606***
	(30.502)	(65.009)	(24.454)
Observations	1089	1084	1084
R^2	0.070	0.124	0.113
Non-PE Mean Dep. Var	1124.733	1081.431	1081.431

Standard errors in parentheses

Note: standard errors clustered at cluster level. All regressions include field office area fixed effects. MHT p-value uses Jones et al. (2018) implementation of Westfall and Young (1993) correction for multiple hypothesis testing.

* $p < .1$, ** $p < .05$, *** $p < .01$

If practices matter, or drive the effects of the program, we might also expect farmers who already adopt improved practices to benefit less from the program than farmers who are less well-informed. Table 7 carries out an analysis with our primary sample to address this question. While the data on improved practices is less detailed in our baseline data than in the follow-up, we have a few ways of measuring improved practices. Column (1) of Table 7 interacts treatment with a dummy for whether or not the farmer already used an improved seed variety at baseline. This interaction term is negative and significantly different from zero. Columns (2) and (3) carry out a similar analysis for

Table 7: Yields per acre and baseline practice use, primary sample

	(1)	(2)	(3)	(4)	(5)
1AF participant=1	305.691***	243.044***	179.004***	489.591**	490.440**
	(88.237)	(65.424)	(60.765)	(219.881)	(220.952)
Used Improved Seed on Maize or Beans=1	48.304				
	(54.617)				
bl_useddap=1		91.505*			
		(54.216)			
bl_usedcan=1			44.253		
			(58.692)		
bl_practices=1				167.078*	
				(95.239)	
bl_practiceindex=1					106.860
					(98.923)
bl_practiceindex=2					184.598*
					(98.229)
bl_practiceindex=3					180.593*
					(102.264)
1AF participant=1 × Used Improved Seed on Maize or Beans=1	-171.346*				
	(92.847)				
1AF participant=1 × bl_useddap=1		-107.687			
		(80.661)			
1AF participant=1 × bl_usedcan=1			-13.037		
			(73.038)		
1AF participant=1 × bl_practices=1				-339.600	
				(223.770)	
1AF participant=1 × bl_practiceindex=1					-299.327
					(230.273)
1AF participant=1 × bl_practiceindex=2					-342.436
					(228.417)
1AF participant=1 × bl_practiceindex=3					-357.053
					(230.436)
Married	98.795	102.415	96.242	98.593	102.309
	(62.622)	(63.673)	(62.903)	(63.164)	(63.942)
Household size	5.824	5.051	5.367	5.797	5.761
	(7.242)	(7.081)	(7.101)	(7.156)	(7.221)
Father, 2ary school (0/1)	8.992	4.052	4.832	7.703	4.650
	(44.995)	(45.708)	(45.089)	(45.093)	(44.989)
Farm labor >50% income	28.731	28.569	29.776	28.404	29.264
	(48.797)	(50.949)	(49.902)	(49.530)	(50.657)
Observations	719	719	719	719	719
R^2	0.111	0.110	0.107	0.111	0.113

Standard errors in parentheses

Note: standard errors clustered at cluster level. The outcome variable is maize yields per acre. All results are from OLS regressions which include field office area fixed effects.

* $p < .1$, ** $p < .05$, *** $p < .01$

whether or not the farmer used DAP or CAN, and while the interaction terms are negative. they are not statistically significant.

Columns (4) and (5) explore instead an index indicating the number recommended practices that a farmer adopted at baseline. This index is different from the one discussed in Section 4 as we don't have information on aspects like row spacing at baseline. The first of the two columns employs a dummy variable equal to one if the farmer adopted one or more recommended practices at baseline, and the second includes an indicator variable for each level of the index. Again, all the interaction terms are negative but not statistically different from zero.

Taken together, these results provide some evidence that farmers learn from the 1AF program— i.e., that information was indeed a constraint— and that they continue using improved practices to some extent, but that these improved practices do not translate into substantially improved yields. Table 7 suggests that the farmers who had access to improved seeds at baseline reap less benefits from the program. Foreshadowing the analysis in the next section, this perhaps hints that access—perhaps to improved inputs or to credit to purchase those inputs—play an important role.

5.2 Credit

The simplest test of whether credit matters maybe to examine whether or not treatment impacts vary by baseline credit access. In other words, do farmers for whom some of the key constraints that the program targets were less binding at baseline appear to reap lower benefits from the treatment? For example, if the credit component of the program were the key to its success, we could imagine that farmers who already had access to credit at baseline may see less of a boost from participating.

5.2.1 Direct analysis of credit constraints

We have in our data two measures of credit access: a narrower definition that includes only respondents who reported having received a loan at baseline ("reported loan use"), and a broader definition that also includes respondents who report that they *could* get a loan if needed ("credit access").[6] It is important to note here that these questions were not targeted to input loans, and we have reason to believe that cash loans may work quite differently from input loans. Setting aside

[6]The "reported loan use" question asked respondents "Have you received a loan or loans formally or informally other than from OAF in the past year?" Credit access is elicited by asking respondents who said they had *not* received a non-1AF loan in the past year, "...would [you] have been able to obtain a loan in the past year?"

this concern for now, we interact these rudimentary measures of credit access at baseline with the treatment. Table 8 reports the results of this simple analysis on per-acre maize yields.

Table 8: Yields per acre and credit

	Primary Sample		Full Sample	
	(1)	(2)	(3)	(4)
	maize_pa	maize_pa	maize_pa	maize_pa
1AF participant=1	111.867*	153.981***	189.570***	168.853***
	(66.075)	(44.927)	(45.784)	(29.697)
Credit access 2016=1	-106.178**		-4.079	
	(44.929)		(31.024)	
1AF participant=1 × Credit access 2016=1	78.417		-34.585	
	(71.758)		(47.886)	
Received loan at baseline=1		-87.235**		41.030
		(43.516)		(33.493)
1AF participant=1 × Received loan at baseline=1		53.088		-19.905
		(72.025)		(44.410)
Past 1AF participant=1			-38.044	-39.316
			(24.582)	(24.434)
Married	106.730*	100.086	53.654	50.809
	(61.473)	(62.556)	(34.026)	(33.993)
Father, 2ary school (0/1)	-2.668	1.673	1.368	0.020
	(44.776)	(44.973)	(24.883)	(25.075)
Farm labor >50% income	22.122	23.226	46.158*	47.560*
	(50.801)	(50.612)	(24.628)	(24.549)
Prev. 1AF knowledge	38.990	43.421	72.655***	72.028***
	(63.067)	(61.983)	(26.301)	(26.436)
Intercropped 2016	-12.794	-16.679	3.224	1.286
	(45.153)	(45.487)	(24.876)	(24.884)
Maize acres, 2016	34.448	34.489	13.500	11.699
	(27.630)	(27.369)	(15.278)	(15.452)
Household size	4.211	3.835	2.066	2.037
	(7.652)	(7.631)	(4.333)	(4.338)
Observations	719	719	1999	1999
R^2	0.114	0.112	0.085	0.085
Non-PE. Ctrl Mean Dep. Var				

Standard errors in parentheses

Note: standard errors clustered at cluster level. All results are from OLS regressions which include field office area fixed effects.

* $p < .1$, ** $p < .05$, *** $p < .01$

Columns (1) and (2) of Table 8 show results in the smaller, primary sample and columns (3) and (4) show the same regressions including the pre-exposed farmers. The latter two columns should be interpreted with some caution since the pre-exposed farmers' credit access may have been affected by their previous interactions with 1AF (if their previous participation was in 2016), making credit access and loan receipt bad controls in those regression. Focusing therefore on the first two columns, we can see that the treatment effects are very consistent with the main analysis in Section 4, but that farmers with baseline credit access seem to have lower yields for both measures of credit access. The interaction terms between baseline credit access and 1AF treatment status is positive but not significantly different from zero. This suggests that at least using this imperfect measure of credit access, treatment effects do not vary starkly with baseline credit access.

5.2.2 Indirect analysis of credit constraints: land assets

Given our concerns about the baseline credit variables, we look at the credit question from a few additional angles: first, we examine whether treatment effects vary by collateralizable assets, i.e., do treatment effects vary with baseline land access? Second, we turn to our pre-exposed sample and examine whether or not these farmers, having internalized the knowledge from their past 1AF training, are able to obtain yields approaching those of treatment farmers. If not, this could suggest that credit constraints may prevent pre-exposed farmers from obtaining the higher yields that they reaped while they participated in the program. Third, we examine currently-treated farmers' behaviors and yields on the land that they are farming but that they did not enroll in 1AF. If behaviors differ on treated and untreated plots, this can provide suggestive evidence weighing the relative importance of credit constraints versus information.

Figure 4 shows results using Generalized Random Forest methods. Covariates include all pre-specified variables plus an index of soil quality on farmers' fields; all other covariates are fixed at the median. In future work, we will complement the GRF analysis with the recently developed machine learning methods in Chernozhukov et al. (2018) as they appear to be more appropriate in settings with many covariates. The results in Figure 4 suggest that the program is most effective for farmers with the smaller land assets at baseline. Given 1AFs focus on small farmers, this is perhaps unsurprising, but since farmers with low levels of land assets are also likely to have worse access to input credit, it may be suggestive that credit constraints matter.

Table 9 shows a parametric version of the analysis in Figure 4: we regress maize yields in 2017 on the treatment dummy, as well as an interaction with baseline acres planted in maize. Columns (1) and (3) show results interacting continuous maize acreage, while columns (2) and (4) use an indicator variable for planting more than three-quarters of an acre of maize. The continuous interactions

Table 9: Yields per acre and baseline acres cultivated

	Primary Sample		Full Sample	
	(1)	(2)	(3)	(4)
	maize_pa	maize_pa	maize_pa	maize_pa
1AF participant=1	186.569***	215.515***	248.140***	229.293***
	(62.845)	(54.736)	(37.882)	(34.372)
1AF participant=1 × Maize acres, 2016	-20.065		-81.600***	
	(49.454)		(29.070)	
.75+ acres last season		8.750		106.281***
		(56.889)		(35.989)
1AF participant=1 × .75+ acres last season		-96.513		-126.666***
		(73.662)		(44.786)
Past 1AF participant=1			-35.855	-36.091
			(25.054)	(24.813)
Married	99.751	102.287	50.694	46.960
	(63.397)	(62.793)	(34.028)	(33.964)
Father, 2ary school (0/1)	-1.892	-0.883	0.690	0.773
	(45.197)	(45.239)	(24.739)	(24.363)
Farm labor >50% income	27.152	28.714	43.549*	43.966*
	(50.321)	(50.475)	(24.227)	(24.218)
Prev. 1AF knowledge	38.733	39.427	70.467***	69.447***
	(62.905)	(62.007)	(26.308)	(26.222)
Intercropped 2016	-17.508	-17.980	1.534	0.715
	(45.689)	(45.530)	(24.904)	(24.968)
Maize acres, 2016	40.724	46.019	50.688**	-5.839
	(36.061)	(32.713)	(20.791)	(19.389)
Maize acres, 2016			0.000	
			(.)	
Maize acres, 2016	0.000			
	(.)			
Household size	3.509	3.418	2.094	1.656
	(7.640)	(7.510)	(4.289)	(4.228)
Observations	719	719	1999	1999
R^2	0.109	0.112	0.089	0.090
Non-PE. Ctrl Mean Dep. Var				

Standard errors in parentheses

Note: standard errors clustered at cluster level. All results are from OLS regressions which include field office area fixed effects.

* $p < .1$, ** $p < .05$, *** $p < .01$

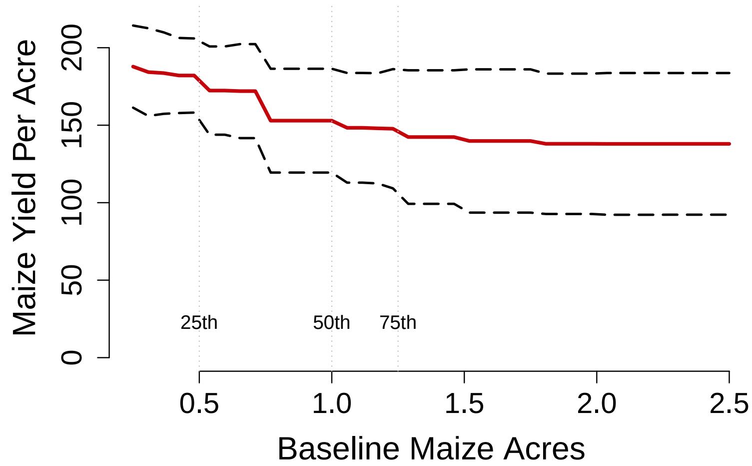

Figure 4: Generalized Random Forest (GRF) analysis: Predicted treatment effects over the support of baseline maize acres (other covariates held at median)

are negative, but only statistically different from zero in the full sample where we might worry about past 1AF participation of some of the sample affecting how many acres farmers choose. Note, however, that the analysis in Section 4 may assuage concerns along these lines, as it did not reveal any impact of program participation on the number of acres planted in maize.

Similarly, the binary interactions have the signs that we would expect based on the GRF analysis, but are only significant in the full sample. Figure 5 show the average marginal effect of treatment for farmers above and below the 0.75-acre threshold at baseline, and the associated 95% confidence intervals. These analyses provides some evidence that treatment has a larger impact on farmers with less land assets. While land assets at baseline may be a correlated with credit access, we still want to caution the reader that this remains speculative as we do not have randomized evidence of the effect of extending input loans to farmers with and without credit access.

5.2.3 Indirect analysis of credit constraints: fertilizer use

Another way that we can examine credit constraints is to return to input use. Section 5.1 evaluated whether or not input use defined as an either-or decision seemed relevant, but we have additional data on the usage intensity that we can exploit. Figures 1 and 2 shown in Section 2 showed that the unexposed control and treatment farmers look very similar in terms of the amount of fertilizer that

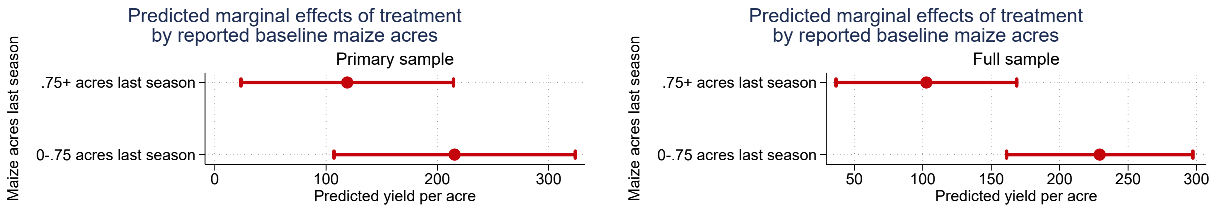

Figure 5: Marginal effect of treatment by baseline maize acres

they apply on their fields at baseline (the red and green lines). If we additionally include pre-exposed farmers **(Figure 6; the blue and purple densities)**, we can observe that the pre-exposed are not only more likely to use fertilizer in general, they also use substantially more of both DAP and CAN than do control farmers. This figure is somewhat difficult to interpret, since the pre-exposed farmers could actually be enrolled in the program during the baseline season, which could increase their fertilizer rates mechanically. It could also be because these self-selected farmers were already using more fertilizer before ever enrolling.

Turning to the current season, we can also examine whether the fertilizer amounts used by pre-exposed and unexposed farmers varies once they are treated. Figure 7 splits the analysis by pre-enrollment status and divides a farmer's fields into enrolled and non-enrolled plots (blue and purple lines represent non-enrolled plots). We can see that the impact on fertilizer use is completely coming from the plots that are currently enrolled in the program. This holds for pre-exposed farmers and new farmers alike. Given that pre-exposed farmers presumably appreciated the program (after all, they chose to re-enroll!) why is there not more spillover onto their non-enrolled plots? It appears that the purple distributions, representing pre-exposed treated farmers' non-enrolled plots, lie slightly to the right of the new treated farmers' intensity distribution, but we cannot reject that these are drawn from the same distribution.

Perhaps the most compelling explanation for the lack of spillovers between enrolled and non-enrolled plots is that of credit constraints. If farmers can only afford to purchase the fertilizer amounts that the program provides them with on loan, then they may be unable to afford to bring

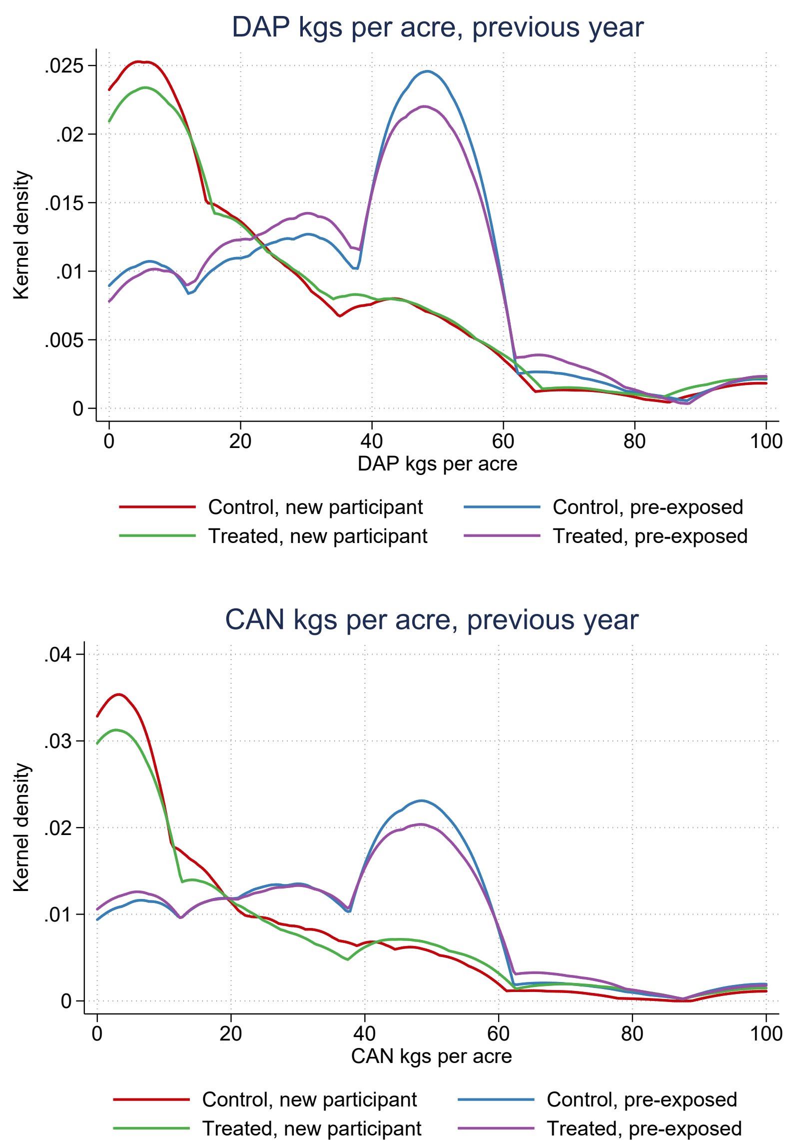

Figure 6: Fertilizer intensity, baseline

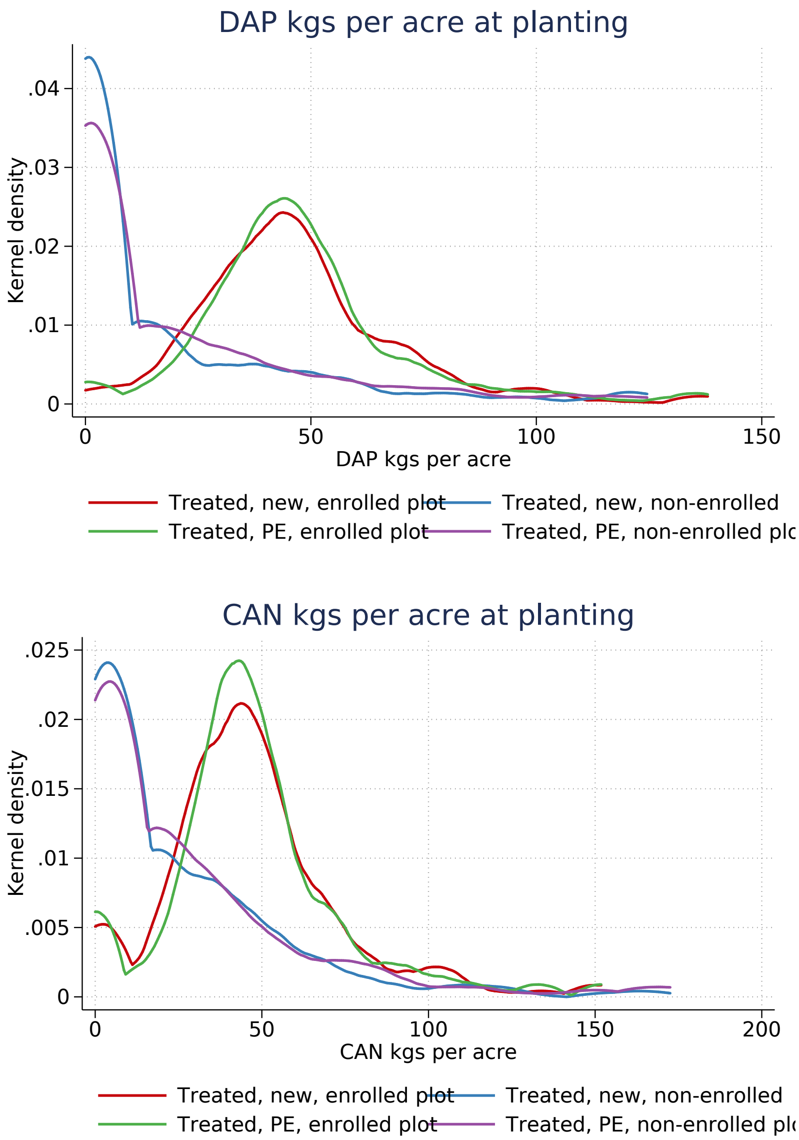

Figure 7: Fertilizer intensity, current season

their other fields up to the recommended fertilizer intensity levels. Granted, alternative explanations abound: farmers may choose to enroll only the plot that they believe will have the highest returns to the program. We have little ability to examine this hypothesis directly. Yet another explanation could be that the inputs that 1AF provides are of higher quality, and farmers do not want to spend money on inferior inputs. As Bold et al. (2017); Tjernström et al. (2018) have demonstrated, fertilizer is often missing advertised nutrient levels. As a next step on this project, we will explore simulating counterfactual yields under full nutrient content to dig further into this question.

While we are still unable to separate the various program components in an entirely satisfactory manner, we believe that the results from the last two sections provide reasonably strong evidence that relaxing the information constraint on its own unlikely suffices. Work in progress includes examining matching farmers in treatment and control on fertilizer intensity to examine whether we can estimate participation effects beyond increased input use.

6 Robustness checks

In this section, we explore the robustness of our primary results to different samples, specifications, and types of inference. Table 10 replicates the analysis in Table 3, except using the full sample that includes pre-exposed farmers. In columns (1) and (2), we can see that the per-acre and per-farmer maize yield impacts are very similar to those using the primary sample—while the point estimates are smaller, they are not statistically different from the results reported in Section 4. Similarly, the results shown in column (3) shows that treatment impacts on farm profits are a bit smaller than those using the primary sample, but not by very much. Furthermore, all of the results remain statistically greater than zero and economically meaningful even when using the full sample. This suggests that including the pre-exposed farmers does not dramatically bias the estimated treatment effects.

To assure the reader that the inclusion of covariates in our preferred specifications is not biasing our results, Table 11 replicates the analysis from Table 3 including only the treatment dummy (columns 1, 3, and 5) as well as regressions excluding covariates but including site fixed effects (columns 2, 4, and 6). The stability of these results across the various specifications should help persuade skeptical readers that covariate adjustment is not driving our results.

Table 10: Primary outcomes, full sample

	(1) Maize PA	(2) Maize Per Farmer	(3) Profit
1AF participant	164.551***	281.794***	45.374**
	(25.917)	(49.697)	(18.523)
Past 1AF participant	-33.953	3.233	2.787
	(25.121)	(63.445)	(22.682)
Married	56.876	40.816	3.902
	(34.417)	(73.664)	(27.238)
Father, 2ary school (0/1)	0.417	203.929***	70.936***
	(24.297)	(51.499)	(19.083)
Farm labor >50% income	36.643	102.909*	41.141*
	(26.277)	(55.483)	(20.923)
Used ag tech 2016	-8.145	-44.779	-22.872
	(32.909)	(67.917)	(25.067)
Prev. 1AF knowledge	52.420*	63.991	16.601
	(28.322)	(55.526)	(21.012)
Intercropped 2016	2.690	-52.260	-12.370
	(25.419)	(58.359)	(21.511)
Credit access 2016	-19.559	-27.312	-19.833
	(24.272)	(52.687)	(19.371)
Maize acres, 2016	25.481	393.505***	129.954***
	(16.498)	(56.672)	(20.521)
Household size	2.400	41.364***	13.601***
	(4.280)	(9.691)	(3.564)
Maize yield/acre, 2016	0.116***	0.379***	0.124***
	(0.032)	(0.085)	(0.031)
FAW Incidence	-4.268*	1.163	0.767
	(2.255)	(4.320)	(1.612)
Observations	1946	1785	1785
R^2	0.097	0.225	0.182
MHT p-value for 1AF	0.000	0.000	0.016
Control Mean Dep. Var	1143.077	1145.997	377.541

Standard errors in parentheses

Note: standard errors clustered at cluster level. All regressions include field office area fixed effects. MHT p-value uses Jones et al. (2018) implementation of Westfall and Young (1993) correction for multiple hypothesis testing.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 11: Primary outcomes, primary sample

	Maize PA		Maize Per Farmer		Profit	
	(1)	(2)	(3)	(4)	(5)	(6)
1AF participant	183.748***	170.206***	376.650***	335.889***	85.175*	70.035**
	(46.300)	(34.661)	(131.786)	(95.411)	(46.110)	(33.776)
Observations	754	754	679	679	679	679
R^2	0.033	0.100	0.028	0.148	0.011	0.126
MHT p-value for 1AF	0.000	0.000	0.005	0.001	0.067	0.040
Control Mean Dep. Var	1124.733	1124.733	1081.431	1081.431	353.263	353.263
Site FE	N	Y	N	Y	N	Y

Standard errors in parentheses

Note: standard errors clustered at cluster level. All regressions include field office area fixed effects. MHT p-value uses Jones et al. (2018) implementation of Westfall and Young (1993) correction for multiple hypothesis testing.

* $p < .1$, ** $p < .05$, *** $p < .01$

To further prod the fidelity of our inference, we take a leaf from Young (2017) and implement a randomization inference procedure using Heß (2017). For each regression, we randomly permute the treatment dummy (maintaining cluster-level treatment assignment) 5,000 times, and for each permutation calculate the hypothetical treatment effect. As with the standard adjustments for multiple hypothesis testing, this method also does not result in any changes to the significance level of the treatment coefficients. Figures 8-9 show the density of post-permutation treatment effect estimates compared to the estimate from the true assignments.

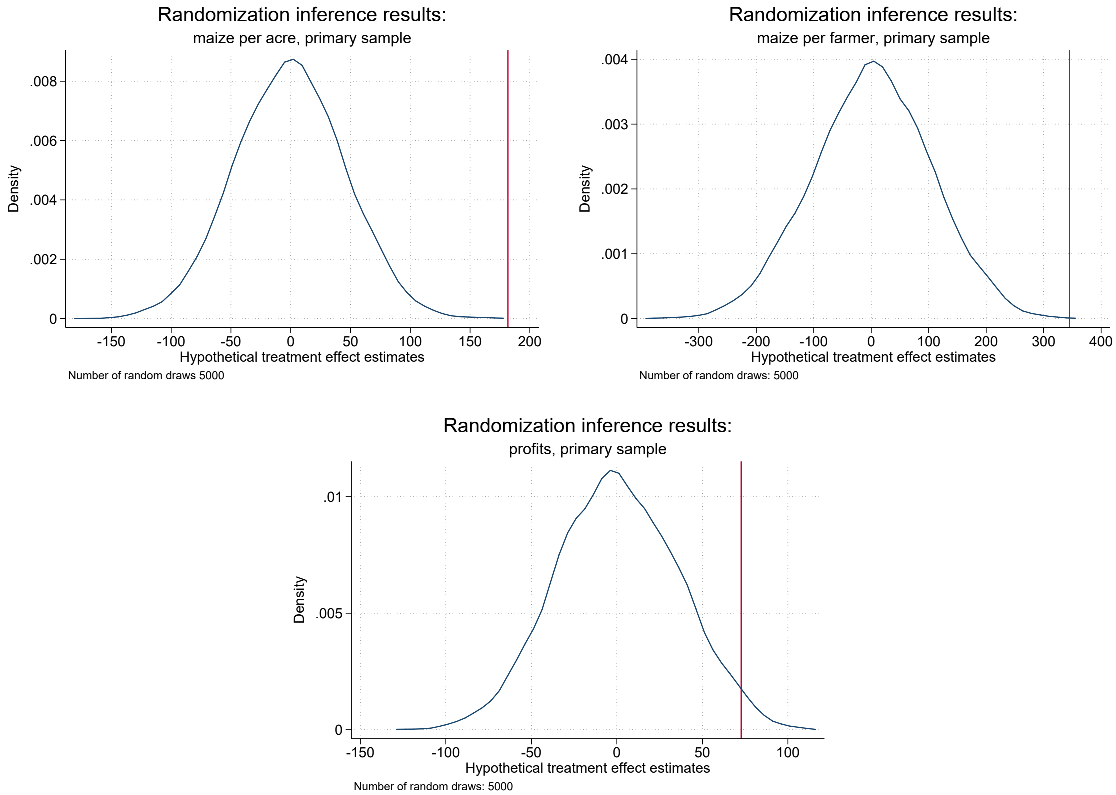

Figure 8: Randomization inference results with primary sample

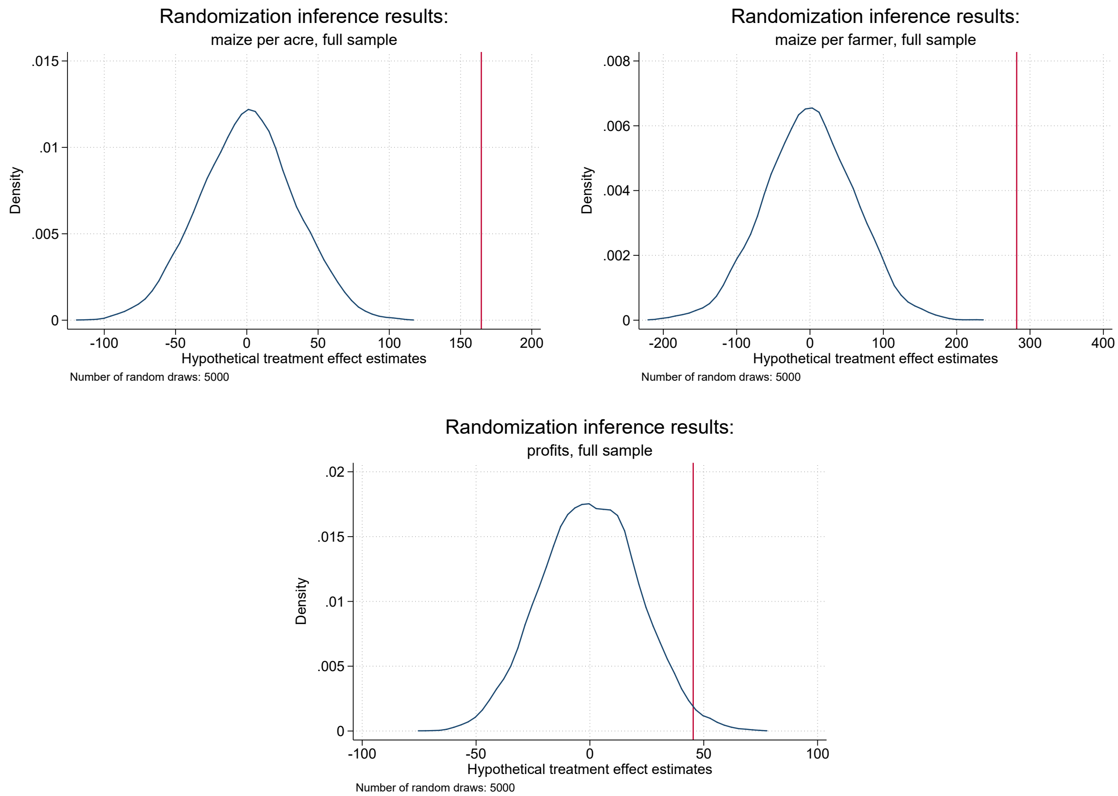

Figure 9: Randomization inference results with full sample

7 Conclusion

This paper presents evidence that a bundled intervention can be effective at increasing smallholder productivity in a region where few interventions have succeeded at this goal. We demonstrate that the effect is robust to a variety of specifications and robustness checks, and try to disentangle some of the mechanisms underlying the program's effectiveness.

In future work, we plan to further explore the mechanisms driving program effects. Linking this work more closely with information on input quality and soil quality could further demonstrate how local supply chains may limit the persistence of the effect of participating. Additionally, in the spirit of Rosenzweig and Udry (2018), we could use regional weather data to explore the inter-temporal external validity of the results. Finally, we have detailed information on labor allocation across different plots, which we hope to use to better understand how farmers allocate complementary inputs.

References

Ashraf, N., X. Giné, and D. Karlan (2009): "Finding Missing Markets (and a Disturbing Epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya," *American Journal of Agricultural Economics*, 91, 973–990.

Bandiera, O., R. Burgess, N. Das, S. Gulesci, I. Rasul, and M. Sulaiman (2017): "Labor markets and poverty in village economies," *The Quarterly Journal of Economics*, 132, 811–870.

Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Pariente, J. Shapiro, B. Thuysbaert, and C. Udry (2015): "A multifaceted program causes lasting progress for the very poor: Evidence from six countries," *Science*, 348, 1260799–1260799.

Beegle, K., L. Christiaensen, A. Dabalen, and I. Gaddis (2016): *Poverty in a Rising Africa*, The World Bank.

Bold, T., K. C. Kaizzi, J. Svensson, and D. Yanagizawa-Drott (2017): "Lemon Technologies and Adoption: Measurement, Theory and Evidence from Agricultural Markets in Uganda*," *The Quarterly Journal of Economics*, 132, 1055–1100.

Carter, M. R., R. Laajaj, and D. Yang (2013): "The Impact of Voucher Coupons on the Uptake of Fertilizer and Improved Seeds: Evidence from a Randomized Trial in Mozambique," *American Journal of Agricultural Economics*, 95, 1345–1351.

——— (2014): "Subsidies and the persistence of technology adoption: Field experimental evidence from Mozambique," Working Paper, NBER.

Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernandez-Val (2018): "Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments," Working Paper, NBER.

Conley, T. G. and C. R. Udry (2010): "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, 100, 35–69.

Deutschmann, J. W. and E. Tjernström (2018): "The impact of One Acre Fund's small farm program," Technical Report.

Dubey, P. and R. N. Yegbemey (2017): "Technical support to the impact evaluation of the Core One Acre Fund Program on Yields and Profits of Maize Farmers in Teso, Kenya," Field Report, International Initiative for Impact Evaluation.

Duflo, E., M. Kremer, and J. Robinson (2008): "How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya," *American Economic Review: Papers and Proceedings*, 98, 482–488.

Elabed, G. and M. Carter (2014): "Ex-ante impacts of agricultural insurance: Evidence from a field experiment in Mali," Working paper.

Fairbairn, A., H. Michelson, B. Ellison, A. Maertens, and V. Manyong (2017): "Adverse Selection in Fertilizer Markets: Evidence from Tanzania," Working Paper.

Feder, G., R. E. Just, and D. Zilberman (1985): "Adoption of Agricultural Innovations in Developing Countries: A Survey," *Economic Development and Cultural Change*, 33, 255–298, articleType: research-article / Full publication date: Jan., 1985 / Copyright © 1985 The University of Chicago Press.

Giné, X. and D. Yang (2009): "Insurance, credit, and technology adoption: Field experimental evidence from Malawi," *Journal of Development Economics*, 89, 1–11.

Gollin, D., C. W. Hansen, and A. Wingender (2018): "Two blades of grass: The impact of the green revolution," Working Paper, NBER.

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

Hess, S. H. (2017): "Randomization inference with Stata: A guide and software," *Stata Journal*, 17, 630–651.

Intermedia Development Consultants (2017): "Review of One Acre Fund Data Collection: Step 3 Report," Field Report.

Jones, D., D. Molitor, and J. Reif (2018): "What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study," Working Paper, NBER.

Karlan, D., E. Kutsoati, M. McMillan, and C. Udry (2011): "Crop Price Indemnified Loans for Farmers: A Pilot Experiment in Rural Ghana," *Journal of Risk and Insurance*, 78, 37–55.

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry (2014): "Agricultural Decisions after Relaxing Credit and Risk Constraints," *The Quarterly Journal of Economics*, 129, 597–652.

Kirimi, L., N. Sitko, T. S. Jayne, F. Karin, M. Muyanga, M. Sheahan, J. Flock, and G. Bor (2011): "A farm gate-to-consumer value chain analysis of Kenya's maize marketing System," Technical Report WPS 44/2011, Tegemeo Institute of Agricultural Policy and Development.

Ligon, E. and E. Sadoulet (2008): "Estimating the effects of aggregate agricultural growth on the distribution of expenditures," Background Paper, World Bank, Washington, D.C.

Maertens, A., H. Michelson, and V. Nourani (2017): "How do Farmers Learn from Extension Services? Evidence from Malawi," Working Paper.

Magruder, J. R. (2018): "An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries," *Annual Review of Resource Economics.*

Mobarak, A. and M. Rosenzweig (2012): "Selling formal insurance to the informally insured," Working Paper.

of Kenya, G. (2010): "Agricultural Sector Development Strategy 2010-2020," Technical Report.

Rosenzweig, M. and C. Udry (2018): "External Validity in a Stochastic World: Evidence from Low-Income Countries," Working Paper.

Sheahan, M., R. Black, and T. S. Jayne (2013): "Are Kenyan farmers under-utilizing fertilizer? Implications for input intensification strategies and research," *Food Policy*, 41, 39–52.

Tjernström, E., M. R. Carter, and T. Lybbert (2018): "The dirt on dirt: Soil characteristics and variable fertilizer returns in Kenyan maize systems," Working Paper.

Westfall, P. H. and S. S. Young (1993): *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*, John Wiley & Sons, Inc.

Young, A. (2017): "Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," Working Paper.

Appendix

A Primary analysis: additional information

Maize yield measurements were taken by enumerators on two 40-square-meter areas selected from farmer plots. Enumerators marked each area before farmers harvested any maize. For treated farmers, two areas were marked each on enrolled and non-enrolled land. Additionally, if farmers planned to harvest green maize from some parts of their plot, enumerators marked areas within these parts for measurement. Yield variables use the average dry weight from the two marked areas. For treated farmers, we average yields across enrolled and non-enrolled land, weighted by the proportion of land farmers enrolled.

We calculate projected revenues using average market prices from nearby vendors covering post-harvest months, multiplied by 1.08 to account for typical price increases over the consumption/selling season. We calculate farmer costs using program costs and self-reported costs for treated farmers, and self-reported seed and fertilizer costs for control farmers. Farm labor is valued using a local day wage for agricultural labor, devalued by 50% (roughly the rural unemployment rate according to DHS data).