Relaxing multiple agricultural productivity constraints at scale

Joshua W. Deutschmann   Maya Duru
Kim Siegal   Emilia Tjernström

October 2022

Abstract

No single constraint can explain the stagnant agricultural productivity growth in sub-Saharan Africa. Most interventions that relax individual barriers to productivity have delivered disappointing results. We evaluate an at-scale program that targets several productivity constraints with a bundled intervention, using a randomized controlled trial in western Kenya. Program participation increases maize yields by 26%, total maize output by 24%, and profits by 17%. While we cannot directly test whether the program's success is due to its bundled nature, we find patterns in the data that are consistent with this hypothesis.

1 Introduction

How should scarce public resources be allocated across sectors of the economy to bolster economic growth and development? This question occupies a central role in development economics and policy debates. As a large contributor to GDP, employment, and food security, the agricultural sector is often at the heart of such discussions. A growing body of evidence demonstrates that agricultural productivity plays a key role in the economic development process by reducing poverty (see de Janvry and Sadoulet 2010 for a review), increasing national welfare (Bravo-Ortega and Lederman, 2005; Gollin, Hansen, and Wingender, 2021; Ligon and Sadoulet, 2011; Ravallion and Chen, 2007), and driving structural change (Jayne, Chamberlin, and Benfica, 2018; McArthur and McCord, 2017).

Nonetheless, knowing that agriculture drives growth is not enough. An efficient allocation of resources requires answers to many other questions: are investments in agriculture the best use of public funds or could other sectors generate higher returns? Should an agriculture-led growth strategy target smallholder farmers or promote larger-scale, commercial operations? Can existing interventions designed to drive agricultural productivity growth be successful at scale?

We contribute to this policy discussion by answering a set of slightly narrower research questions, which we consider fundamental for answering these bigger-picture policy questions. First, we ask whether a scaled-up agriculture-focused program for African smallholders can generate economically meaningful productivity gains. Having answered this question in the affirmative—a pre-condition for the viability of smallholder agriculture-led growth strategies—we try to understand what distinguishes this program from the many unsuccessful examples in the literature. Specifically, we ask whether the bundled nature of the program, which helps relax multiple constraints that small farmers face, might be part of its success. Further, we examine treatment effect heterogeneity to understand whether the program is more effective for certain types of farmers, and move beyond yields to look at whether the net effects on participant welfare are positive. Finally, we try to answer the broader question of whether the program increases social welfare, by comparing our estimated benefits to the program's costs.

Our primary contribution is to provide experimental evidence of the productivity impacts of an at-scale program run by One Acre Fund (1AF), an established non-governmental organization (NGO). 1AF was founded in 2006 and currently works with over one million farm households in seven countries across Eastern and Southern Africa. We evaluate their "full service" smallholder farmer program, which offers enrolled farmers a bundle that includes loans for improved seeds and fertilizer, training on modern agricultural techniques,

and crop insurance. Our identification strategy relies on a cluster-randomized experiment, in which we randomized access to the program for farmer groups in western Kenya. Despite the agricultural sector's economic and social importance, there is a dearth of rigorous empirical evidence on bundled agricultural extension interventions. Research efforts often suffer from issues of measurement, selection, and comparability across programs (Aker, 2011; Anderson and Feder, 2007); high-quality evaluations of programs operating at scale are especially rare.

In contrast to most of the growing body of experimental evidence in this space, we find that 1AF program participation causes statistically and economically significant increases in productivity and total output, defined as maize output per acre and total maize production at the farm level, respectively. Other studies—including those that detect meaningful effects on farmer practices or input use—rarely detect measurable increases in yields (Cole and Fernando, 2021; Udry, di Battista, Fosu, Goldstein, Gurbuz, Karlan, and Kolavalli, 2019) or profits (Beaman, Karlan, Thuysbaert, and Udry, 2013). In our preferred specification, we find that 1AF program participation increases maize productivity by 26% and total maize output by 24%.

Agricultural productivity in many African countries remains substantially lower than other regions of the world (Block, 2014; World Bank, 2008). Investments in the agricultural sector in many African countries have yielded considerable evidence on what fails and a limited catalog of successes (Suri and Udry, 2022). This has led some to question the viability of agriculture-led growth strategies. Further, opinions differ on whether African governments should direct resources towards the smallholders that farm the majority of the region's arable land, or if developing medium- and large-scale farms is more cost-effective. In recent years, many African governments have shifted resources towards developing medium- and large-scale farms (Jayne, Chamberlin, Traub, Sitko, Muyanga, Yeboah, Anseeuw, Chapoto, Wineman, Nkonde, and Kachule, 2016). Still, smallholder agriculture constitutes a large fraction of the region's total agricultural production and its farmers make up a substantial portion of the population.[1] Our findings suggest that there is scope for optimism about these small-scale producers' growth potential: well-designed programs can increase agricultural productivity among small farmers in sub-Saharan Africa.

Moving beyond agricultural productivity, we seek to understand whether the program increases farmer welfare. Using rich plot-level data on input use and farming practices, we find that program participation changes several different dimensions of farming behavior. We assign a monetary value to the main costly inputs, including labor, fertilizer, and

---

[1]This holds true for many low-income countries: Lowder, Skoet, and Raney (2016) estimate that 72% of farms in the world use less than one hectare of land.

3

purchased seed, and subtract them from the market value of the maize outputs to obtain a measure of farmer profits. On average, we find that program participation is profitable, with the net value of maize production increasing by about 17%. Although we cannot fully capture welfare changes from other income sources, the median farmer in our sample dedicated all of their cultivated acreage to maize at baseline. This suggests that changes in the net value of maize production constitute an important part of agricultural income, and likely overall household income.

This paper also aims to advance our understanding of the obstacles to technology diffusion and productivity growth on smallholder farms.[2] Three constraints stand out in the literature as key barriers to adoption of improved agricultural technologies: credit, risk, and information (Feder, Just, and Zilberman, 1985; Magruder, 2018). 1AF offers farmers a "market bundle" of components that are closely tied to these constraints: participating farmers receive input loans for high-quality seeds and fertilizer, crop insurance, and training on improved farming practices. None of the program ingredients are novel on their own; in fact, they overlap substantially with the kinds of programs that the World Bank promoted as part of their agricultural extension investments in the 1960s and 1970s (Birkhaeuser, Evenson, and Feder, 1991; Chambers, 1983).

These similarities aside, the program that we study distinguishes itself from its antecedents by tightly bundling the component parts into a single unified package (Tinsley and Agapitova, 2018). This kind of bundling is not typical of government-led extension, nor is it common in the economics literature, which tends to evaluate the impact of relaxing a single constraint at a time.[3] A key reason that researchers have shied away from bundled programs is that bundling makes it difficult to disentangle the underlying economic mechanisms. Nonetheless, our data allow us to examine several farming choices that precede productivity and output along the causal chain. The resulting empirical patterns suggest that both information and credit constraints are binding in the study context. More broadly, the results are consistent with the notion that multiple constraints matter for farmers' decisions and outcomes.

The theoretical literature on poverty traps is a notable exception to the frequent focus on single constraints. A growing empirical body of evidence complements this theoretical literature by investigating the possibility that poor households may need large, bundled interventions in order to move out of poverty (Balboni, Bandiera, Burgess, Ghatak, and

---

[2]See Magruder (2018) and Suri and Udry (2022) for recent reviews.

[3]Two recent papers show promising impacts of interventions that target multiple constraints in the context commercially-focused smallholder agriculture (Arouna, Michler, and Lokossou, 2021; Deutschmann, Bernard, and Yameogo, 2021).

Heil, 2022; Bandiera, Burgess, Das, Gulesci, Rasul, and Sulaiman, 2017; Banerjee, Duflo, Goldberg, Karlan, Osei, Pariente, Shapiro, Thuysbaert, and Udry, 2015). Our results echo recent findings from multi-faceted anti-poverty programs, where individual intervention components seem unable to generate the same magnitude of effects.

Another contribution of our study is that we take seriously the broader question of whether the benefits that we have estimated above are worth the overall costs of the 1AF intervention. Specifically, we conduct a cost-benefit analysis to determine the net economic impact of the 1AF program in western Kenya. Rather than defend a particular set of assumptions, we account for different types of uncertainty using Monte Carlo simulations. The simulations produce distributions of the net benefits that emerge under a range of different assumptions. This analysis suggests that the program increases social welfare and that our results are robust to a wide range of scenarios: we obtain a positive net present value in 79-87% of simulations under different scenarios.

Finally, we use a variety of heterogeneity analyses to explore whether the average treatment effects mask underlying impact heterogeneity. Treatment effect heterogeneity could indicate potential efficiency gains from programmatic changes or improved participant targeting. We find little evidence of systematic heterogeneity in treatment effects on productivity and output. There are several potential explanations for this lack of heterogeneity, including the potential screening effect of the small upfront fee that 1AF requires prior to enrollment—another difference between 1AF and its many predecessors.[4]

One interpretation of our results is that past research may have underestimated the potential for well-designed programs to increase agricultural productivity. To the extent that governments and donors make their investment choices based on evidence, these findings could lead to more efficient decision-making. We discuss some candidate explanations for why our findings differ from those in the literature in Section 7.

An obvious limitation of our study is that it focuses on one crop in one agricultural region, suggesting some caution in generalizing our results to other seasons and other populations (Rosenzweig and Udry, 2020). However, our results are consistent with 1AF's non-experimental monitoring and evaluation results, and consistent with farmers' own revealed preference, with the program continuing to grow rapidly and many farmers choosing to re-enroll (Deutschmann and Tjernström, 2018; One Acre Fund, 2020). Taken together, our findings on the broad impacts of the program and its cost-effectiveness suggest there

---

[4]Studies of health products in low-income countries generally find cost-sharing mechanisms to be ineffective screening devices (Ashraf, Berry, and Shapiro, 2010; Cohen and Dupas, 2010; Tarozzi, Mahajan, Blackburn, Kopf, Krishnan, and Yoong, 2014). However, Beaman, Karlan, Thuysbaert, and Udry (2020) find evidence of positive selection into an agricultural credit program in Mali.

is indeed a role for targeting smallholder farmers in an agriculture-led growth strategies. Supporting and learning from effective programs operating at scale will be a key component of any such strategy.

2 Context, data, and experimental design

2.1 One Acre Fund's program in Kenya

We analyze the main operating model of an established agricultural NGO, One Acre Fund. Founded in 2006, the organization has grown rapidly in the last several years: enrollment has grown from 200,000 farm households in 2014 to more than one million in 2020 (One Acre Fund, 2020). 1AF's "full service program" provides farmer groups with loans for improved seeds and fertilizer, regular training on modern agricultural techniques, crop and funeral insurance, and market facilitation support to help farmers obtain higher prices for their output (Tinsley and Agapitova, 2018). In the context we study, input credit is given as joint liability loans to farmer groups, organized by geographical area and typically comprised of 8-12 farmers.

The program's focus on credit, risk, and information has support in the economics literature, which has accumulated substantial evidence that failures in these domains hinder farmers' ability or willingness to adopt improved agricultural technologies (Feder et al., 1985; Magruder, 2018). 1AF's decision to focus on providing high-quality inputs also has empirical support: a growing literature establishes that fertilizer and seeds sold in the region's local markets often fall short of quality standards, and that farmer perceptions of input quality matter even when inputs meet quality standards (Bold, Kaizzi, Svensson, and Yanagizawa-Drott, 2017; Hsu and Wambugu, 2022; Michelson, Fairbairn, Ellison, Maertens, and Manyong, 2021; Tjernström, Carter, and Lybbert, 2018).

Prior to enrolling in the smallholder farmer program, participants sort into self-selected farmer groups. Participants then choose how much of their land to enroll, in increments of 0.25 acres. The agricultural credit and input quantity provided by 1AF is determined as a function of land enrollment. Farmer groups are jointly responsible for repayment of these loans. Loan terms are relatively flexible, allowing repayment at any time during the growing season.[5] Groups must complete repayment in full within a two-week grace period after harvest.[6] To alleviate information constraints, 1AF field officers conduct repeated

---

[5]Field, Pande, Papp, and Rigol (2013) suggest flexibility in loan repayment is important in encouraging illiquid investments among microentrepreneurs.

[6]Historically, repayment rates exceed 97% (Tinsley and Agapitova, 2018).

trainings and provide handouts on the benefits and proper use of improved inputs. To alleviate risk, 1AF provides yield-index insurance based on crop cuts and forgives a portion of the input loan in case of crop failure (Tinsley and Agapitova, 2018).

This experiment was conducted in western Kenya, where 1AF has operated for more than ten years. Agriculture contributes 51% to Kenya's GDP and is dominated by small-scale producers (Government of Kenya, 2010). Despite the importance of the agricultural sector to the economy, most small-scale farmers are not running successful micro-enterprises. Households in Kenya typically derive their income from the production of a variety of crops (Sheahan, Black, and Jayne, 2013) and average productivity often fails to meet households' dietary needs (Kirimi, Sitko, Jayne, Karin, Muyanga, Sheahan, Flock, and Bor, 2011).

Although small-scale farmers in Kenya produce a range of crops across a diverse production environment, 1AF's program focuses on the dominant staple crop: maize. Maize is important both to the economy and for food security. 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 seed varieties and inorganic fertilizer at higher rates than neighboring countries, yields remain low. Further, the country remains a net importer of maize, despite policy targets to the contrary.

2.2 Experimental design

Recruitment, enrollment, and program implementation for this experiment took place in the Teso region of Kenya, following 1AF's standard protocol. Farmers who satisfied the basic program criteria paid a participation deposit of approximately $5 USD.[7] Once participants had self-selected into groups of 8-12 farmers, the randomization was conducted at the level of clusters, which consisted of 2-4 of these joint-liability farmer groups. This aggregation was designed to minimize potential spillovers by maximizing the distance between clusters.

After participants had paid the fee and signed the contract, they were informed of the randomization, which took place as a public lottery. Farmer groups randomly assigned to the control group were informed they would only be able to enroll in the program in the following year. They were offered a roughly 20% discount for enrollment in the following season, as well as a compensation package consisting of a bundle of household goods valued

---

[7]The participation criteria for the standard program is possession of a phone number and national identification. Study participants had to consent to be part of the study and plan to cultivate at least 0.25 acres of maize. After contract signing, 1AF informed farmers about the study, that their participation would be voluntary, and went over informed consent.

at the cost of participation.[8]

Dealing with sample contamination

While the ideal location for this study would arguably be an area where 1AF had never operated, this proved difficult in Kenya. Given 1AF's long presence in the country, all "untouched" regions would have been unrepresentative of the production environments and farmer populations that 1AF usually engages with. As a second-best option, the evaluation team decided to sample villages within district where 1AF was working, but where the program had not yet recruited farmers directly from those villages.

In practice, although 1AF had never offered its program to the sampled villages, about 64% of the farmers in our sample had managed to access it anyway, by "commuting" to neighboring villages to participate. The proportion of such "pre-exposed" farmers does not differ across treatment and control groups. We may nevertheless be concerned that the pre-exposed farmers introduce bias into our results. Accordingly, we report our main results both for the full sample and for a "primary sample," which refers to the sample of farmers who had never participated in 1AF programming.

The sign of any potential bias depends on at least two factors: the persistence of treatment effects and the correlation between selection and program returns. Suppose that 1AF participation changes farmers' production processes for the better. If pre-exposed farmers continue to operate more efficiently in subsequent seasons, then treatment effects estimated using the full sample may be attenuated.

We also do not know the sign of the correlation between self-selection into early access and returns to enrollment. If the most eager farmers have higher returns, we would expect the impacts on our primary sample of hold-out farmers to be a lower bound of the true impacts. If these hold-out farmers instead resemble the never-adopters in Suri (2011) and have surprisingly high returns—but perhaps face high costs of participation—then it is possible that focusing solely on the primary sample might overestimate impacts.

2.3   Data

Baseline data collection took place in November and December of 2016—after initial enrollment but before contract signing. A public lottery assigned clusters of farmer groups to treatment in January 2017. Enumerator teams conducted input use surveys after the planting of the main season in 2017, from April through June. Enumerators collected fresh

---

[8]Households received a bag and thermos with a total value below $15 USD. These items were chosen because they are common in the study area and unlikely to have significant resale value.

weight measurements between July-August 2017 and dry weight measurements between July-September 2017. Enumerators also conducted market surveys throughout the implementation of the project. They collected data on input prices from local agrovets and maize and other output prices in local markets between May and October 2017.

2.3.1 Outcome variables

Our main analysis centers on maize productivity per acre and total maize output. Our measure of productivity per acre captures a direct comparison of average yields on the land treated farmers enroll in the 1AF program to the average yield in the control group. Since farmers rarely enroll all of their land, we might wonder about various forms of spillovers across a given farmer's plots.[9] As such, we also measure and show a comparison between maize yields on treated farmers' non-enrolled land and control farmers. To further characterize the holistic impacts of program enrollment on farmers' agricultural production, net of any substitution between plots, we generate and analyze an estimate of total maize output based on crop cut yields and total land cultivated.

To obtain "crop cut" yields, enumerators collected and physically weighed fresh and dry harvests from two randomly placed 40-square-meter boxes. Cultivated land sizes were measured by GPS readings, with enumerators walking the boundaries of each plot three times.[10] Based on the objective measures of land size and crop cut yields, we generate more standard per-acre productivity and total output measures.

Enumerators collected separate crop cut measurements for enrolled land and non-enrolled land in the treatment group. While our measure of productivity per acre refers to treated farmers' per-acre yield on their enrolled plot, total maize output is a weighted average of per-acre yields on the enrolled vs. non-enrolled plot (with weights proportional to the amount of land in each category). Control farmers did not separate their fields into enrolled and non-enrolled sub-plots. For them, productivity equals per-acre yields on a randomly-selected maize plot. Total output is scaled up to their total land under maize cultivation.[11]

To further quantify the potential welfare effects of the program, we calculate profits

---

[9]For example, there could be positive knowledge spillovers to non-enrolled plots, or participants may reallocate scarce complementary inputs such as labor to the enrolled plot. Conversely, farmers might prefer to spread 1AF-provided inputs across multiple plots in an effort to reduce risk.

[10]A growing body of research documents non-classical measurement error in self-reported land size and harvests in developing-country studies (see e.g. Abay, Bevis, and Barrett 2021; Carletto, Savastano, and Zezza 2013; Desiere and Jolliffe 2018; Gourlay, Kilic, and Lobell 2017). Self-reported harvests are additionally subject to recall bias.

[11]Appendix B provides additional variable construction details.

as the value of output less farmers' costs. Revenues are the product of total output and average market prices from nearby vendors. To understand input use, we elicit the quantity of inputs farmers applied to their plots. For farmers in the treatment group, we further distinguish between inputs applied to enrolled and non-enrolled land. To attribute costs to the inputs that farmers used, we rely on two sources of data. For inputs sourced from 1AF, we rely on administrative data on prices charged which are standard for all farmers. For all other inputs used by farmers in both treatment and control groups, we assign prices based on mean prices measured in a sample of local markets.[12]

Labor costs are challenging to measure in this context, given the prevalence of unpaid labor. To reduce recall bias, we elicit early-season paid and family labor use in a survey administered shortly after planting (including self-reported labor for land preparation, plowing, and planting). Late-season paid and family labor use, including for weeding and harvest, was collected shortly after harvest. We ascribe a monetary value to family (unpaid) labor by calculating the mean day wage reported within the sample. We show two estimates of profit: one in which we value unpaid labor at 50% of the paid day wage (based on rural unemployment rates in Kenyan DHS data) and one in which we value unpaid labor at the full day wage.[13] We then multiply these labor prices by the number of person-days of unpaid labor.

2.3.2 Attrition

Attrition in our data is mostly due to missing variables, not to participants dropping out of the sample entirely. During data cleaning, we are primarily forced to drop observations for missing land size and harvest data.[14] In Appendix C, we include a more complete description of attrition from our sample and implement several imputation strategies to test the robustness of our results. We show that our results are consistent across various robustness checks, including missing-data imputation methods.

---

[12]Overall, these data confirm that 1AF charges prices for seeds and fertilizer that are comparable to local market prices.

[13]Our approach to valuing unpaid labor matches closely with Agness, Baseler, Chassang, Dupas, and Snowberg (2022), which suggests valuing unpaid labor at 60% of the paid labor day wage.

[14]Land size data was collected during and immediately after harvest time, and in some cases enumerators were unable to complete the land survey on some or all plots. This issue affects 142 farmers in our sample, and occurs more often among treatment farmers (7%) than among control group farmers (4%). The increased attrition among treatment group farmers is primarily driven by missing land measurement on non-enrolled land. Harvest data was recorded both immediately after harvest (the "fresh weight") and after drying (the "dry weight"). Our analysis relies on dry weight estimates of yields, which are a more comparable measure across farmers since moisture content before drying can vary. For 56 farmers, we observe fresh weights but not dry weights. For 268 farmers, we are missing both fresh and dry weights. We drop a further 139 observations due to inconsistencies in outcome data or missing control variables.

2.3.3 Sample description and balance tests

We report summary statistics and balance tests on our pre-specified control variables in Table 1, which confirms the importance of agriculture for this population: nearly eighty percent of participants earned more than half of their income from farm labor in the year prior to the study. On average, participants cultivated roughly one acre of maize, harvesting about half a ton per acre. While three-quarters of the sample used improved agricultural technologies at baseline, average input intensity is low. About half of the respondents report some knowledge of 1AF planting practices, which is primarily driven by the pre-exposed farmers.

Table 1: Baseline balance by treatment assignment, full sample

Variable	(1) Control Mean (SE)	(2) Treatment Mean (SE)	Difference (2)-(1)
Married (0/1)	0.88 (0.01)	0.88 (0.01)	0.01
Household head has secondary school (0/1)	0.38 (0.02)	0.44 (0.02)	0.05*
Household income >50% from farm labor (0/1)	0.78 (0.01)	0.78 (0.01)	0.00
Used improved ag technology in 2016 (0/1)	0.78 (0.01)	0.81 (0.01)	0.02
Reports knowledge of 1AF practices (0/1)	0.47 (0.02)	0.51 (0.02)	0.04
Intercropped maize and beans in 2016 (0/1)	0.48 (0.02)	0.48 (0.02)	0.00
Reports having credit access in 2016 (0/1)	0.71 (0.01)	0.74 (0.02)	0.03
Household size	6.66 (0.08)	6.90 (0.09)	0.26**
Acres under maize cultivation in 2016	1.01 (0.03)	1.05 (0.03)	0.04
Maize yield (kg/acre) in 2016	537.01 (13.52)	587.77 (16.25)	51.99**
F-statistic (test of joint significance)			0.82
Number of observations			1896

Note: field office area fixed effects included in all balance test regressions. Standard errors are clustered at the treatment assignment (farmer group cluster) level.

Although household size, education, and baseline maize yields differ significantly across treatment and control, an F-test of joint orthogonality of the variables in Table 1 does not reject the null that all the variables are jointly orthogonal to treatment status. None of the three significantly different variables are highly correlated with our outcome variables, so the differences are unlikely to affect our results.[15] Nevertheless, we report results both with and without these control variables.

In addition to balance checks for treatment and control groups, we compare baseline characteristics of our two sub-populations: the "primary" sample (farmers who have never

---

[15]It may seem counter-intuitive that baseline maize yields do not correlate with maize yields, but these variables are very noisy and low autocorrelation is common.

enrolled in 1AF) and the "pre-exposed," who had previously enrolled in 1AF programming. The two groups together make up the "full sample." These tables are shown in Appendix A. We observe imbalance in some individual variables but fail to reject the null that the variables are jointly orthogonal to treatment status.

We also compare participants in the primary sample to those who enrolled prior to the study (Table A.3). At baseline, farmers who self-selected into early program access are more educated, cultivate more maize land, report using improved inputs and having access to credit, and are more likely to have prior knowledge of 1AF practices. We cannot determine whether this is due to early enrollees being better off to start or whether program participation contributed to the observed differences.

2.3.4 Data collection and quality validation

The data collection was directly managed by the NGO, which has an independent data collection department. To assuage potential concerns about the independence of the research design, the research team took several steps. First, the International Initiative for Impact Evaluation (3ie) helped design and review all parts of the trial (the experimental design, field protocols, sampling, randomization, and data collection instruments (Dubey and Yegbemey, 2017)). Second, the data collection followed best-practice protocols for data collection, including back-checks at all phases of data collection. Third, 1AF contracted an independent firm, Intermedia Development Consultants, to carry out a three-step audit of the data collection. Fourth, as described above, maize yields were physically weighed and cultivated maize acreage was measured with GPS, respectively, which should minimize the potential for social desirability bias. All weighing and surveying was carried out by enumerators hired by the 1AF Monitoring and Evaluation department. These were not program staff, were from outside the study area , and did not know the sample farmers. Finally, two of the authors on this paper were brought in as independent evaluators. This part of the team reviewed the pre-analysis plan (PAP) prior to data collection and independently conducted the data cleaning, variable construction, and analysis (Deutschmann and Tjernström, 2018).

3 Empirical strategy

We obtain intent-to-treat estimates of program impacts by estimating the following regression with OLS:

$$y_{ics} = \alpha + \beta T_{cs} + [\delta X_{ics}] + \gamma_s + \varepsilon_{ics} \tag{1}$$

where y_{ics} denotes the outcome for individual i in cluster c and strata s, with strata defined based on 1AF field-office areas. The indicator variable T_{cs} takes on a value of one for clusters assigned to treatment. X_{ics} is a vector of controls, γ_s is a strata fixed effect to account for variations in the probability of assignment to treatment, and ε_{ics} is clustered at the level of treatment assignment (farmer group clusters).

In the body of the paper, we report results which include a control for past exposure to the 1AF program where appropriate, but no other controls. We show results in Appendix C which include either our pre-specified set of controls[16] or a set of controls selected via a double lasso procedure (Belloni, Chernozhukov, and Hansen, 2014; Urminsky, Hansen, and Chernozhukov, 2016). The standard errors in our main results are not adjusted for multiple hypothesis testing. The interpretation of our results does not change when we account for multiple hypothesis testing using Westfall and Young (1993) and Holm-Bonferroni methods, nor when we apply Fisher's exact test in a randomization inference procedure following Young (2019).

4   Results: productivity and output

The primary goal of the 1AF small farmer program is to increase the maize productivity of participating farmers. We consider two related outcomes in this section: per-acre maize productivity and total maize production. Productivity per acre is particularly important if we are concerned about allocative efficiency and structural transformation (Dercon and Gollin, 2014; Gollin, 2015). Understanding total maize production allows us to understand how the program affects overall farmer welfare, a topic we discuss more extensively below in Section 6.

Participation in the 1AF program has an economically and statistically significant effect on both per-acre productivity and total maize production. In columns 1 and 2 of Table 2, we find that maize productivity (kg per acre) increases by 26-27% in the treatment group. As described above, for treated farmers we consider per-acre productivity on the land farmers enrolled in the 1AF program. On average, enrolled land among the treatment group represents 63% of total cultivated maize acreage. We also observe maize production on the non-enrolled land in the treatment group. In columns 3 and 4, we show that farmers

---

[16]Our PAP specified X_{ics} in Eq. 1 as including controls for household demographics (marital status, household size, education, credit access, land ownership, agricultural reliance) and baseline agricultural characteristics (technology use, intercropping, knowledge of 1AF practices). Additionally, the PAP proposed to include a spillover inverse probability weight, with spillover likelihood measured by a farmer's total agricultural contacts that were randomized into treatment. Including this weighting does not change the statistical significance or qualitative interpretation of our results (results available upon request).

in the treatment group are not increasing per-acre productivity on their enrolled land at the expense of productivity on their non-enrolled land. Per-acre productivity on non-enrolled land is statistically similar to productivity among control farmers.

Table 2: Productivity and output

	Productivity per acre enrolled vs. control		Productivity per acre non-enrolled vs. control		Total output per farmer	
	(1)	(2)	(3)	(4)	(5)	(6)
1AF participant	295.43***	306.82***	-52.15	-37.81	264.90***	278.78***
	(37.230)	(37.970)	(36.790)	(41.090)	(85.730)	(90.560)
Pre-exposed		20.05		24.28		192.43**
		(34.130)		(33.870)		(74.320)
1AF participant		-14.05		-10.32		-33.33
× pre-exposed		(49.290)		(55.560)		(120.710)
1AF + (1AF × Pre-exposed)		292.77***		-48.13		245.45***
		(30.37)		(40.47)		(66.92)
Control group mean	1128.39	1150.49	1128.39	1150.49	1082.38	1164.00
Observations	682	1896	614	1701	682	1896

Note: This table presents results from linear regressions of the outcomes at the top of each column on the treatment dummy. Productivity per acre is measured as kgs of maize per acre cultivated, and total maize output is measured in kgs. In columns 1 and 2, enrolled refers to the land farmers in the treatment group enrolled in the 1AF program, which may be less than the full amount of land on which they cultivate maize. In columns 3 and 4, we include only farmers who also cultivated maize on land they did not enroll with 1AF and compare the yields from that non-enrolled land to yields from control group farmers. Standard errors (in parentheses) are clustered at the treatment assignment (farmer group cluster) level. Columns 1, 3, and 5 include only farmers who had never previously participated in the 1AF program. Columns 2, 4, and 6 additionally include a sample of farmers who had previously enrolled in the 1AF program (see Section 2.2). All regressions include strata (field office) fixed effects.

Columns 5 and 6 of Table 2 corroborate that farmers in the treatment group are not only more productive per acre, but that they increase overall maize production by 24% relative to control farmers. Farmers are using their land more efficiently and harvesting more maize as a result.

The effects on productivity and overall production are not statistically different between pre-exposed farmers in the treatment and control groups. The point estimates on the interaction between random assignment to the program and past participation are negative, suggesting the benefits may be slightly lower for those farmers, but the coefficients are imprecisely estimated. Participation in the 1AF program seems to be beneficial for farmers beyond a single season, a finding which is perhaps unsurprising from a revealed preference perspective given the large number of farmers who re-enroll in the program every year.

5 Mechanisms

The bundled nature of the intervention we study in this paper makes it difficult to fully disentangle the mechanisms driving the effects we observe. Previous work studying farmer constraints to technology adoption and productivity has typically focused on isolating single constraints. We provide suggestive evidence consistent with multiple constraints playing a role in farmers' decisions and maize production, suggesting the bundled program is effective at shifting farmers' constraints along multiple dimensions. Indeed, it is highly likely that there are important complementarities between these dimensions, and the effects of the bundled program may be greater than the sum of its parts.

5.1 Information and behavioral changes

We first examine several practice choices farmers make during production. These are primarily or purely informational constraints: farmers could adopt these practices even absent the 1AF program or the credit it provides for input purchases. If we observe treatment effects on adoption of these practices, it would be consistent with the program easing information constraints and inducing farmers to adopt improved practices.

Table 3 reports treatment effects on key practices, estimated with linear probability models. The outcome variable in columns 1 and 2 is a dummy equal to one if farmers planted their rows within 5 cm of 1AF's recommended spacing. Similarly, columns 3 and 4 is a dummy equal to one if farmers spaced their plants within rows within 5cm of 1AF's recommended spacing. Enumerators measured this spacing at several points on farmers' fields, and we compare the average of those measurements to recommendations from 1AF training guides. In columns 5 and 6, fertilizer timing equals one if the participant applied fertilizer at the appropriate time of the season. High maize productivity requires substantial amounts of nitrogen—but the timing matters. Maize plants require relatively little nutrient input at the time of planting and are most responsive to inputs later in the season. Accordingly, farmers are recommended to apply DAP (diammonium phosphate) at planting and CAN (calcium ammonium nitrate) several weeks later.

Columns (1)-(4) show that treated participants are more likely than control farmers to follow spacing recommendations. Adherence to recommended spacing increases by almost 60% (160%) for row spacing (plant spacing) among farmers unacquainted with the program. Similarly, treated farmers are 170% more likely to use appropriate fertilizer timing. In column (4), we can see that the treatment did not affect the likelihood that a farmer used an ox-plow.

Table 3: Behavioral changes

	Row Spacing		Plant Spacing		Fertilizer Timing	
	(1)	(2)	(3)	(4)	(5)	(6)
1AF participant	0.22***	0.24***	0.20***	0.21***	0.66***	0.65***
	(0.040)	(0.040)	(0.030)	(0.030)	(0.030)	(0.030)
Pre-exposed		0.07**		0.02		0.15***
		(0.030)		(0.020)		(0.030)
1AF participant		-0.11**		-0.07*		-0.18***
× pre-exposed		(0.050)		(0.040)		(0.040)
1AF + (1AF × Pre-exposed)		0.13***		0.14***		0.47***
		(0.03)		(0.02)		(0.03)
Control group mean	0.37	0.41	0.09	0.13	0.26	0.37
Observations	682	1896	682	1896	682	1896

Note: This table presents results from linear regressions of the outcome variables in each column on the treatment dummy. Row and plant spacing are indicator variables for whether or not a farmer was within 5cm of the recommended spacing, for the rows in which they planted and the plants within rows, respectively. Fertilizer timing is an indicator variable for whether or not farmers applied DAP at planting and CAN at top dressing (the timing recommended by the NGO). Standard errors (in parentheses) are clustered at the treatment assignment (farmer group cluster) level. Columns 1, 3, and 5 include only farmers who had never previously participated in the 1AF program (the "primary" sample). Columns 2, 4, and 6 additionally include a sample of farmers who had previously enrolled in the 1AF program (the "full" sample; see Section 2.2). All regressions include strata (field office) fixed effects.

It is not surprising that, absent the program, experienced farmers may have overlooked things like the proper timing of fertilizer, which might seem like basic steps to improve productivity. Research suggests that even highly-experienced producers can fail to notice crucial features of the production process (Hanna, Mullainathan, and Schwartzstein, 2014) or less-salient profitability margins (Beaman et al., 2013; Duflo, Kremer, and Robinson, 2008). Fertilizer recommendations in the region typically focus on the amount of fertilizer rather than on application timing.[17]

Table 3 also reports estimated coefficients on previous program exposure and its interaction with the treatment. The results show that the treatment effects on 1AF-promoted practices are substantially larger for new participants (85% for row spacing, 50% for plant spacing, and 38% for fertilizer timing). Nevertheless, pre-exposed farmers also experience positive and significant treatment effects, ranging from 13% (14%) for row (plant) spacing to 47% for fertilizer timing. It appears that even seemingly basic information about practices may bear repeating multiple times, albeit with diminishing returns over time.

5.2   Liquidity constraints and investment

Input intensification is an important contributor to agricultural productivity gains. Table 4 therefore examines program impacts on the intensive margins of production. The main variables are expenditure on fertilizer, seeds, labor (both paid and unpaid, with the latter valued as described in section 2.3), and land. Enrolled farmers spend 75% more on fertilizer, 26% more on seeds, 18% more on paid labor, 37% more on unpaid labor, and dedicate 12% more land to maize.[18] These are substantial expenditure increases, which we account for in our analysis of profit below.[19]

Relative to results above suggesting diminishing treatment effects on information among pre-exposed farmers, we find no differential treatment effect on investment among that group. Pre-exposed farmers in the treatment group increase their investment along all the dimensions we consider by a comparable magnitude to the primary sample. If information were the primary constraint preventing farmers from using appropriate quantities of seeds and fertilizer, we might expect the treatment effect on these margins to decrease

---

[17]While extension manuals often distinguish between fertilizer application at planting and at top dressing, they rarely emphasize the importance of timing (e.g. National Farmers Information Services 2019).

[18]Farmers in treated and control groups report using a similar fraction of total available land for maize cultivation, with the median farmer reporting 83% of total cultivated acreage is dedicated to maize production.

[19]1AF's average gross margin on inputs is 32%, which is similar to markups in the agro-dealer sector in the region (Tinsley and Agapitova, 2018). Cost increases are therefore unlikely driven by 1AF's input prices.

Table 4: Use of costly inputs

	Fertilizer		Seeds		Paid Labor		Unpaid Labor		Maize Acres	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1AF participant	18.74***	18.98***	3.84***	4.01***	5.32*	5.85*	4.90***	5.11***	0.12*	0.12**
	(1.950)	(1.930)	(0.810)	(0.820)	(2.920)	(3.030)	(1.000)	(1.080)	(0.060)	(0.060)
Pre-exposed		5.62***		1.23		6.17***		-0.96		0.14***
		(1.740)		(0.750)		(1.900)		(0.970)		(0.050)
1AF participant		0.45		0.62		2.89		1.31		-0.03
× pre-exposed		(2.630)		(1.140)		(3.700)		(1.280)		(0.080)
1AF + (1AF × Pre-exposed)		19.44***		4.63***		8.74***		6.41***		0.10**
		(1.51)		(0.71)		(2.15)		(0.83)		(0.04)
Control group mean	20.55	25.23	14.77	15.59	29.63	31.90	14.65	13.69	0.93	0.98
Observations	682	1896	682	1896	682	1896	682	1896	682	1896

Note: This table presents results from linear regressions of the outcome variables in each column on the treatment dummy. Costs are expressed in USD. For more on how we define labor costs, see Appendix B. Columns 1-8 show input costs in USD. Columns 9 and 10 show all acres used for maize cultivation, including both land enrolled in the 1AF program and non-enrolled land. Standard errors (in parentheses) are clustered at the treatment assignment (farmer group cluster) level. Odd columns include only farmers who had never previously participated in the 1AF program. Even columns additionally include a sample of farmers who had previously enrolled in the 1AF program (see Section 2.2). All regressions include strata (field office) fixed effects.

over time as farmers learn, similar to the effects we observe in Table 3. If farmers were marginally credit constrained, we might also expect the credit from the 1AF program to permit increased investments, increased profits, and subsequently similarly reduced effects on these dimensions. We interpret these results as instead suggesting that farmers are substantially credit constrained, and the credit and potential profit from a marginal additional year of participation in the program does not fully resolve the constraints they face.

As further evidence of credit constraints playing a key role in farmers' decisions, we consider major categories of household expenditures reported by farmers before and after treatment assignment. Note that these all cover the period before farmers received any inputs from 1AF. Changes in expenditures here represent responses to an upcoming relaxation of credit constraints. In Table 5, we show that farmers in the treated group increase expenditures on school fees by 50-70% relative to the control group following treatment assignment. This suggests that farmers may have avoided incurring additional debt in the form of school fee loans, or alternatively that they could allocate limited savings towards school fees. We see no detectable changes in medical, celebration, or livestock expenditures. School fees constitute a relatively large fraction of reported household expenditures compared to these other categories, so it may be that a relaxation of credit constraints is relatively less important for those expenditure categories.

Table 5: Major expenditures, 3 months after treatment assignment

	Education		Medical		Celebration		Livestock	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1AF participant	55.78***	54.97**	-0.57	-0.78	3.23	3.60	-1.08	-0.63
	(18.990)	(21.950)	(0.500)	(0.610)	(3.850)	(4.460)	(3.110)	(3.060)
Pre-exposed		56.33***		0.31		0.23		2.90
		(17.090)		(1.160)		(1.530)		(2.450)
1AF participant		-15.77		1.23		-3.84		-2.01
× pre-exposed		(31.680)		(1.770)		(4.590)		(3.740)
1AF + (1AF × Pre-exposed)		39.21**		0.45		-0.24		-2.64
		(19.03)		(1.61)		(1.71)		(2.17)
Control group mean	79.48	110.22	1.09	1.17	1.34	1.64	7.39	8.17
Observations	682	1896	682	1896	682	1896	682	1896

Note: This table presents results from linear regressions of the outcome variables in each column on the treatment dummy. The outcomes in this table show the respondent's expected expenditures for the three months after treatment assignment. Expenditures are in USD. Each regression controls for reported expenditures. in the same category for the three months before treatment assignment. Standard errors (in parentheses) are clustered at the treatment assignment (farmer group cluster) level. Odd columns include only farmers who had never previously participated in the 1AF program. Even columns additionally include a sample of farmers who had previously enrolled in the 1AF program (see Section 2.2). All regressions include strata (field office) fixed effects.

5.3 Heterogeneity

Heterogeneity analysis may allow us to further explore the potential mechanisms driving the average effects we observe above. For example, we observe baseline indicators of self-reported credit access. If program effects were greater for farmers without credit access, we might conclude that the credit component of the program is particularly important. To explore this in an agnostic way about which covariates may be predictive, we implement machine learning methods developed by Chernozhukov, Demirer, Duflo, and Fernandez-Val (2018).

Figure 1 shows the estimated Group Average Treatment Effects that result from applying the methods of Chernozhukov et al. (2018). We observe that treatment effects on both productivity and total maize output show little evidence of heterogeneity between groups, although Figure 1b suggests that effects on total output may be attenuated for some farmers. The estimation procedure also produces what Chernozhukov et al. (2018) term the 'heterogeneity loading' parameter of the Best Linear Predictor of the conditional average treatment effect. For all outcomes and methods, we fail to reject the null hypothesis that this heterogeneity parameter is statistically different from zero.

To further corroborate these results, we conduct a Monte Carlo simulation exercise of the type suggested in Appendix E of Heckman, Smith, and Clements (1997). To simulate the distribution of impact standard deviations under the null hypothesis of no heterogene-

ity, we repeatedly sample the control group to generate synthetic treatment and control groups. This gives us a distribution of the standard deviation of percentile effect differences under the null, which we then compare to the impact standard deviation seen in the data. For all outcomes, we fail to reject the null of no heterogeneity, suggesting that we are unable to detect treatment effect heterogeneity under the assumption of perfect positive dependence between treatment and control outcome percentiles (also called the location shift assumption). Since this is a strong assumption, we also implement the rank preservation test proposed in Bitler, Gelbach, and Hoynes (2005). This tests for rank reversal in baseline characteristics between quartiles of the treatment and control distribution. For each sample definition and outcome variable, we fail to reject the null in the test for joint-orthogonality.[20] These exercises together suggest limited scope to find major heterogeneity in treatment effects. In other words, the program seems broadly effective at relieving multiple constraints for farmers, and there does not seem to be one group which is relatively unaffected by this bundled intervention.
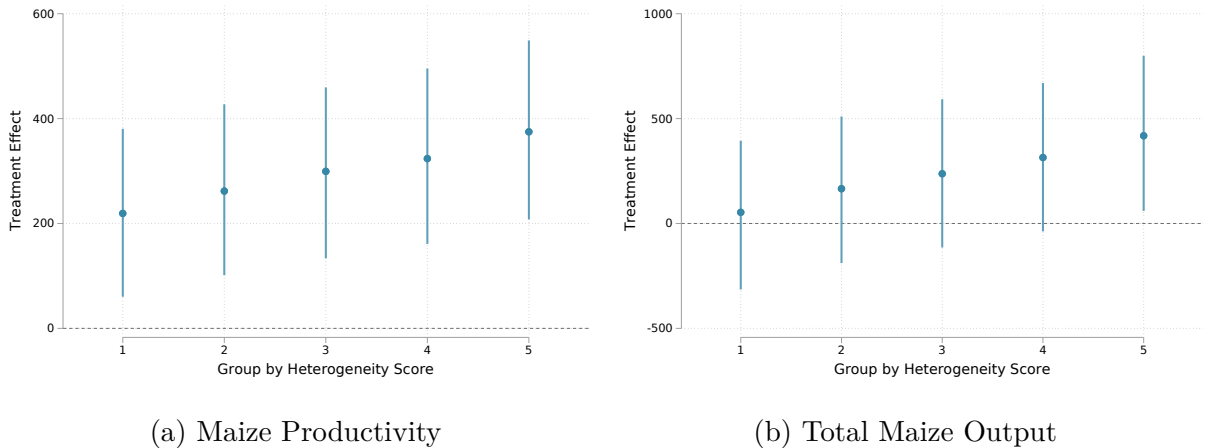

(a) Maize Productivity

(b) Total Maize Output

Figure 1: Sorted Group Average Treatment Effects, estimated with Neural Nets

6   Welfare and Net Benefits

We have shown that participation in the 1AF program increases farmers' maize production, but that participating farmers also increase their expenditures on seeds, fertilizer, and labor. To determine the net effect on farmer welfare, we construct an estimate of farmer maize profits that accounts for the market value of maize production. In Table 6, we estimate that farmer profits in the treatment group increase by $53-61 USD, or 15-17%. This finding

---

[20]Results of both exercises are available from the authors upon request.

is robust to a variety of assumptions about the value of unpaid labor. Our pre-analysis plan specified valuing unpaid labor at 50% of the day wage in local markets. Recent work by Agness et al. (2022) in a similar setting suggests valuing unpaid labor at 60% of the market wage. Even conservatively assigning the full day wage to unpaid work results in qualitatively similar conclusions about profits.

The effect of program participation on profits is attenuated among pre-exposed farmers in the treatment group. Farmers in the pre-exposed group are farming more profitably in general. We cannot say with certainty whether increased profits among pre-exposed farmers represent past effects of program participation or selection into the program in the past. However, given our observations above that some benefits of participation are attenuated among the pre-exposed sample, it is plausible that for some of these farmers the benefits of participating in the full program no longer outweigh the costs. Farmers may nevertheless find it worthwhile given credit constraints and the quality of inputs available through the 1AF program.

Table 6: Profits

	Profit (PAP Definition)		Profit (Full Labor Costs)	
	(1)	(2)	(3)	(4)
1AF participant	58.02*	61.26*	53.12*	56.15*
	(30.200)	(32.490)	(30.020)	(32.190)
Pre-exposed		61.50**		62.46**
		(27.480)		(27.100)
1AF participant		-20.20		-21.51
× pre-exposed		(43.160)		(42.750)
1AF + (1AF × Pre-exposed)		41.06*		34.64
		(24.21)		(23.96)
Control group mean	335.35	364.61	320.70	350.92
Observations	682	1896	682	1896

Note: This table presents results from linear regressions of the outcomes shown at the top of each column on the treatment dummy. Profits are calculated as the value of maize production less costs (see Appendix B), and are reported in USD. Columns 1 and 2 follow the PAP and account for unpaid labor costs by valuing unpaid time at 50% of the average day-wage reported in the sample for paid labor. Columns 3 and 4 instead value unpaid labor at the full average day-wage reported in the sample. Standard errors (in parentheses) are clustered at the treatment assignment (farmer group cluster) level. Odd columns include only farmers who had never previously participated in the 1AF program. Even columns additionally include a sample of farmers who had previously enrolled in the 1AF program (see Section 2.2). All regressions include strata (field office) fixed effects.

6.1 Net benefits under different assumptions

Beyond the direct effect on farmer profits of program participation, we also consider the broader question of whether the benefits that we have estimated above are worth the overall costs of 1AF's intervention. Specifically, we conduct a simple cost-benefit analysis to determine the net economic impact of the 1AF program in Kenya. In international aid discussions, the concept of value for money is often equated with return on investment (ROI) metrics. This approach more closely resembles a financial analysis than an economic cost-benefit analysis; in particular, the latter aims to capture the net contribution that an intervention makes to social welfare.[21]

Rather than rely on single treatment effect point estimates, we carry out Monte Carlo simulations to account for different types of uncertainty. Our main sources of uncertainty are parameter uncertainty (due to the statistical uncertainty around our estimated treatment effects) and what the cost-benefit literature calls *structural uncertainty* (due to the assumptions made during analysis). We already explore the robustness of our main results to several sources of structural uncertainty in Appendix C, but here we look instead at the effect on net benefits.

Table 7: Monte Carlo simulation parameters and probability distributions

Parameter	Distribution	Source
Donor subsidy	$\mathcal{U}(25.3, 5.9)$	1AF administrative data, 2016-2021[a]
Profit, new farmers	$\mathcal{N}(61.3, 32.5)$	Table 6, column (2)[b]
Profit, pre-exposed farmer	$\mathcal{N}(41.1, 32.5)$	Table 6, column (2)[b]

Note:
[a] We obtained internal data from 1AF on the average donor subsidy per enrolled farmer in Kenya for the period 2016-2021. While we cannot share the exact numbers, we show here the mean and standard deviation across the period.
[b] For simplicity, we assume that the distributions of profit treatment effects have the same standard deviation for new and pre-exposed farmers. We conservatively use the larger regression standard error as our estimate of the standard deviation.

The standard approach in economics would be to compute the discounted stream of expected utility from the program and subtract the discounted stream of costs, to arrive at the net present value (NPV) of the program. Instead of utility, we use our treatment effect estimate on profits as the main benefit. Ideally, we would prefer to have consumption data instead of profits, as it more closely maps into utility; however, consumption data was not collected because the primary focus of this study was on maize yields and agricultural

---

[21]To illustrate the difference, a financial analysis does not consider the influence of market distortions on project impacts or costs.

profits. Given imperfect credit markets in this context, profits may be a reasonable proxy for consumption. The result in Table 5, which shows that program participants spend around $55 more on education—a number similar in magnitude to the average profit treatment effect—lends further support to our assumption that participants consume most of the increased income.

Table 7 details the probability distributions that we assume for this probabilistic analysis, and Figure 2 shows the net benefit distributions that emerge from the simulation exercise. The costs of administering the program vary some amount from year to year. To account for this, we use administrative data on the average per-farmer donor subsidy from 1AF's Kenya program spanning the time period 2016-2021. We approximate the variation using a uniform distribution since the year-to-year variation is relatively small and since we don't have strong reasons to believe that any of the data points are more likely than others.

Figure 2: Distribution of net benefits under different assumptions
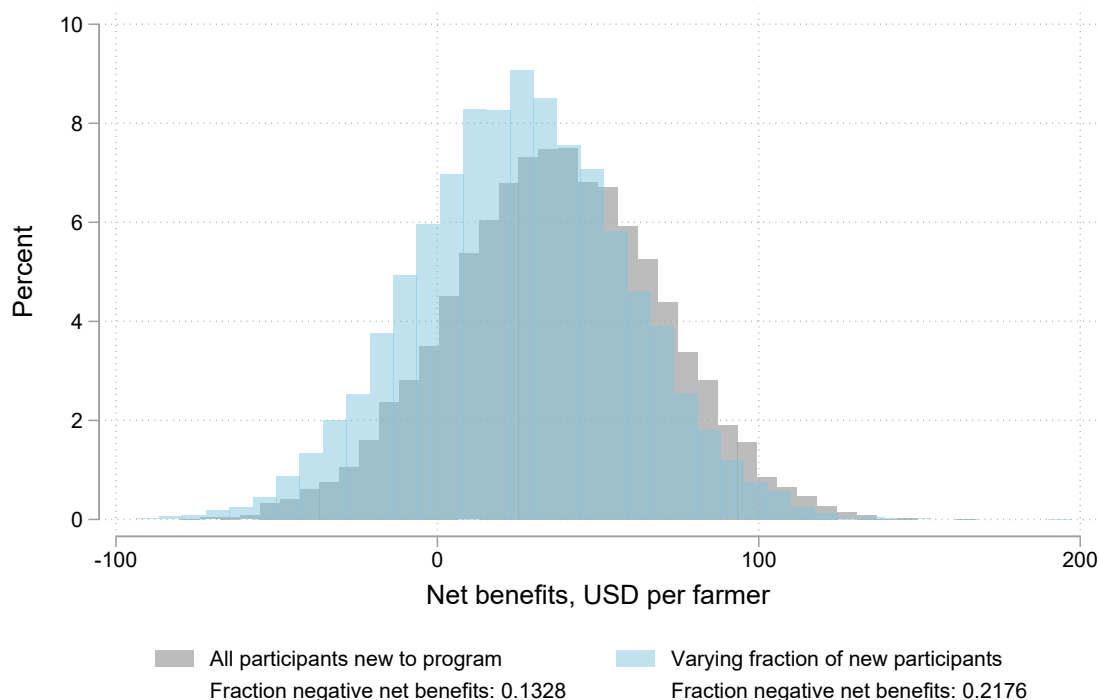
Note: This figure shows the distribution of net benefits under 2 scenarios. The gray distribution shows estimated net benefits when we assume all participants are new to the program. The blue distribution instead draws a random fraction of 'pre-exposed' farmers using the point estimate of profits for this group (see Table 7).

We first look at the scenario where all participants are new to the program. Drawing

per-farmer donor subsidies and average profits across 10,000 trials, we derive the distribution of net benefits shown in Figure 2 by a gray histogram. In this scenario, we estimate that the average net benefits of the program are positive in about 87 percent of cases.

To evaluate the net benefits of a program operating at scale, it may be more plausible to posit that some participants will already have been exposed to the program—much like in our study area. We estimate a second scenario in which we additionally draw a random fraction of 'pre-exposed' respondents, who receive a smaller average profit increase from program participation (shown in the third row of Table 7). The light blue histogram in Figure 2 shows the resulting distribution of estimated net benefits. The median of the distribution is slightly lower, but we nevertheless estimate that the net benefits of the program are positive in about 78 percent of cases.

In summary, our simulations suggest that the net benefit of the program is positive—not only for participants themselves, but taking into account the fraction of program fees that are not covered already by farmer fees. Moreover, given that we find no evidence of heterogeneous treatment effects above in Section 5.3, we expect that the net benefit of participation should be broadly positive for participants. Perhaps most convincingly, 1AF continues to grow rapidly, and we observe that a majority of treated farmers in our sample choose to re-enroll in the program in the subsequent season, and on average they increase the percentage of their land enrolled in the program.

7 Conclusion

We uncover large ATEs from participation in 1AF's core program. The treatment effects are remarkably stable across sample definitions and specifications and we find limited evidence of heterogeneous effects. Our results are consistent with the large theoretical literature on poverty traps as well as empirical work showing that poor households benefit from bundled interventions (Balboni et al., 2022; Bandiera et al., 2017; Banerjee et al., 2015). If we believe that farmers face multiple simultaneous constraints, an intervention that only provides credit or information may not suffice to raise yields and profits in a meaningful way.

While we cannot answer the question of whether a bundled approach is more cost effective than a simpler or more targeted program, our work suggests some considerations that future work on this topic could explore further. In contexts where the nature and severity of market failures vary sharply across individuals or space, it may be prohibitively costly for an organization to target interventions to each locality. A standardized program

that targets multiple constraints may allow an organization to scale across space without tweaking program specifics in each new context. The cost-effectiveness of bundled programs therefore hinges on the costs associated with "over-bundling" relative to the cost of the market research required to tailor programs to each new context.

Returning to the broader question of the role of the agricultural sector in Africa's future and the contribution of smallholder farmers to economic growth, we recognize the limits of a single study in a particular region in a specific year. A deeper discussion of external validity (even as it applies to extrapolating to other 1AF locations) would be valuable, but is beyond the scope of the current paper. Nevertheless, the 1AF model addresses several of the concerns about smallholder farming systems that Collier and Dercon (2014) raise.

A primary reason why small farmers may struggle to compete relates to how scale influences capital costs, logistical capacity, and bargaining power. Large organizations can and do leverage grant funding for working capital from both agro-input suppliers and banks (Tinsley and Agapitova, 2018). Most financial institutions shy away from input loans, preferring to provide credit over shorter time periods. By extending their credit to smallholders, organizations like 1AF can effectively reduce the transaction costs and asymmetric information facing small farmers. 1AF also imports their own inputs, conducts their own quality controls, and manages their own storage facilities. In so doing, they have successfully integrated several parts of the value chain, allowing them to benefit from economies of scale—the benefits of which they can pass on to their clients. In addition, a large organization like 1AF can leverage their scale and human capital to process and distill large amounts of information on agricultural productivity, risks, and complementary practices to share with small farmers when and where they need it. This allows small farmers to benefit from skills and technology transfers otherwise often out of reach for all but the largest market participants.

Perhaps this new evidence will nudge a few cynics into reconsidering the future for (smallholder) agriculture as low-income countries grow and develop. For those already optimistic about the sector, we hope that it provides compelling input into a discussion about optimal instruments for boosting productivity. Thinking about large-scale organizations like 1AF as enabling creative vertically-integrated opportunities for farmers may hold a clue to the types of investments needed to transform African agriculture into a more dynamic sector.

References

ABAY, K. A., L. E. M. BEVIS, AND C. B. BARRETT (2021): "Measurement Error Mechanisms Matter: Agricultural Intensification with Farmer Misperceptions and Misreporting," *American Journal of Agricultural Economics*, 103, 498–522.

AGNESS, D., T. BASELER, S. CHASSANG, P. DUPAS, AND E. SNOWBERG (2022): "Valuing the Time of the Self-Employed," Working Paper 29752, National Bureau of Economic Research.

AKER, J. C. (2011): "Dial "A" for Agriculture: A Review of Information and Communication Technologies for Agricultural Extension in Developing Countries," *Agricultural Economics*, 42, 631–647.

ANDERSON, J. R. AND G. FEDER (2007): "Chapter 44 Agricultural Extension," in *Handbook of Agricultural Economics*, ed. by R. Evenson and P. Pingali, Elsevier, vol. 3 of *Agricultural Development: Farmers, Farm Production and Farm Markets*, 2343–2378.

AROUNA, A., J. D. MICHLER, AND J. C. LOKOSSOU (2021): "Contract Farming and Rural Transformation: Evidence from a Field Experiment in Benin," *Journal of Development Economics*, 151, 102626.

ASHRAF, N., J. BERRY, AND J. M. SHAPIRO (2010): "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia," *American Economic Review*, 100, 2383–2413.

BALBONI, C., O. BANDIERA, R. BURGESS, M. GHATAK, AND A. HEIL (2022): "Why Do People Stay Poor?*," *The Quarterly Journal of Economics*, 137, 785–844.

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.

BEAMAN, L., D. KARLAN, B. THUYSBAERT, AND C. UDRY (2013): "Profitability of Fertilizer: Experimental Evidence from Female Rice Farmers in Mali," *American Economic Review: Papers and Proceedings*, 103, 381–86.

——— (2020): "Selection into Credit Markets: Evidence from Agriculture in Mali," Working Paper.

BELLONI, A., V. CHERNOZHUKOV, AND C. HANSEN (2014): "Inference on Treatment Effects after Selection among High-Dimensional Controls," *The Review of Economic Studies*, 81, 608–650.

BIRKHAEUSER, D., R. E. EVENSON, AND G. FEDER (1991): "The Economic Impact of Agricultural Extension: A Review," *Economic Development and Cultural Change*, 39, 607–650.

BITLER, M. P., J. B. GELBACH, AND H. W. HOYNES (2005): "Distributional Impacts of the Self-Sufficiency Project," Working Paper 11626, National Bureau of Economic Research.

BLOCK, S. (2014): "The Decline and Rise of Agricultural Productivity in Sub-Saharan Africa since 1961," in *African Successes*, University of Chicago Press, for the National Bureau of Economic Research, 13–67.

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.

BRAVO-ORTEGA, C. AND D. LEDERMAN (2005): "Agriculture and National Welfare around the World: Causality and International Heterogeneity since 1960," Working paper 3499, The World Bank.

BURKE, M., L. F. BERGQUIST, AND E. MIGUEL (2019): "Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets," *The Quarterly Journal of Economics*, 134, 785–842.

CARLETTO, C., S. SAVASTANO, AND A. ZEZZA (2013): "Fact or Artifact: The Impact of Measurement Errors on the Farm Size–Productivity Relationship," *Journal of Development Economics*, 103, 254–261.

CARTER, M. R., E. TJERNSTRÖM, AND P. TOLEDO (2019): "Heterogeneous Impact Dynamics of a Rural Business Development Program in Nicaragua," *Journal of Development Economics*, 138, 77–98.

CHAMBERS, R. (1983): *Rural Development: Putting the Last First*, Routledge.

CHERNOZHUKOV, V., M. DEMIRER, E. DUFLO, AND I. FERNANDEZ-VAL (2018): "Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments," Working Paper 24678, National Bureau of Economic Research.

CHERNOZHUKOV, V., I. FERNÁNDEZ-VAL, AND B. MELLY (2013): "Inference on Counterfactual Distributions," *Econometrica*, 81, 2205–2268.

COHEN, J. AND P. DUPAS (2010): "Free Distribution or Cost Sharing: Evidence from a Randomized Malaria Prevention Experiment," *The Quarterly Journal of Economics*, 125, 1–45.

COLE, S. A. AND A. N. FERNANDO (2021): "'Mobile'Izing Agricultural Advice Technology Adoption Diffusion and Sustainability," *The Economic Journal*, 131, 192–219.

COLLIER, P. AND S. DERCON (2014): "African Agriculture in 50 Years: Smallholders in a Rapidly Changing World?" *World Development*, 63, 92–101.

DE JANVRY, A. AND E. SADOULET (2010): "Agricultural Growth and Poverty Reduction," *World Bank Research Observer*, 25, 1–20.

DERCON, S. AND D. GOLLIN (2014): "Agriculture in African Development: Theories and Strategies," *Annual Review of Resource Economics*, 6, 471–492.

DESIERE, S. AND D. JOLLIFFE (2018): "Land Productivity and Plot Size: Is Measurement Error Driving the Inverse Relationship?" *Journal of Development Economics*, 130, 84–98.

DEUTSCHMANN, J. W., T. BERNARD, AND O. YAMEOGO (2021): "Contracting and Quality Upgrading: Evidence from an Experiment in Senegal," Working Paper.

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.

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.

FIELD, E., R. PANDE, J. PAPP, AND N. RIGOL (2013): "Does the Classic Microfinance Model Discourage Entrepreneurship among the Poor? Experimental Evidence from India," *American Economic Review*, 103, 2196–2226.

GOLLIN, D. (2015): "Agriculture as an Engine of Growth and Poverty Reduction," in *Economic Growth and Poverty Reduction in Sub-Saharan Africa*, ed. by A. McKay and E. Thorbecke, Oxford University Press, 91–121.

GOLLIN, D., C. W. HANSEN, AND A. M. WINGENDER (2021): "Two Blades of Grass: The Impact of the Green Revolution," *Journal of Political Economy*, 129, 2344–2384.

GOURLAY, S., T. KILIC, AND D. LOBELL (2017): "Could the Debate Be Over? Errors in Farmer-Reported Production and Their Implications for the Inverse Scale-Productivity Relationship in Uganda," Working paper 8192, The World Bank.

GOVERNMENT OF KENYA (2010): "Agricultural Sector Development Strategy 2010-2020," Technical Report.

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.

HECKMAN, J. J., J. SMITH, AND N. CLEMENTS (1997): "Making The Most Out Of Programme Evaluations and Social Experiments: Accounting For Heterogeneity in Programme Impacts," *The Review of Economic Studies*, 64, 487–535.

HSU, E. AND A. WAMBUGU (2022): "Can Informed Buyers Improve Goods Quality? Experimental Evidence from Crop Seeds," Working Paper.

JAYNE, T., J. CHAMBERLIN, L. TRAUB, N. SITKO, M. MUYANGA, F. K. YEBOAH, W. ANSEEUW, A. CHAPOTO, A. WINEMAN, C. NKONDE, AND R. KACHULE (2016): "Africa's Changing Farm Size Distribution Patterns: The Rise of Medium-Scale Farms," *Agricultural Economics*, 47, 197–214.

JAYNE, T. S., J. CHAMBERLIN, AND R. BENFICA (2018): "Africa's Unfolding Economic Transformation," *The Journal of Development Studies*, 54, 777–787.

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. A. and E. Sadoulet (2011): "Estimating the Effects of Aggregate Agricultural Growth on the Distribution of Expenditures," Working paper 1115, Department of Agricultural & Resource Economics, UC Berkeley.

Lowder, S. K., J. Skoet, and T. Raney (2016): "The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide," *World Development*, 87, 16–29.

Magruder, J. R. (2018): "An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries," *Annual Review of Resource Economics*, 10, 299–316.

McArthur, J. W. and G. C. McCord (2017): "Fertilizing Growth: Agricultural Inputs and Their Effects in Economic Development," *Journal of Development Economics*, 127, 133–152.

Michelson, H., A. Fairbairn, B. Ellison, A. Maertens, and V. Manyong (2021): "Misperceived Quality: Fertilizer in Tanzania," *Journal of Development Economics*, 148, 102579.

National Farmers Information Services (2019): "Field Management – NAFIS," Report, http://www.nafis.go.ke/agriculture/maize/field-management-practices/.

One Acre Fund (2020): "How We Grow," Report, https://oneacrefund.org/what-we-do/how-we-grow/.

Ravallion, M. and S. Chen (2007): "China's (Uneven) Progress against Poverty," *Journal of Development Economics*, 82, 1–42.

Rosenzweig, M. R. and C. Udry (2020): "External Validity in a Stochastic World: Evidence from Low-Income Countries," *The Review of Economic Studies*, 87, 343–381.

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.

SURI, T. (2011): "Selection and Comparative Advantage in Technology Adoption," *Econometrica*, 79, 159–209.

SURI, T. AND C. UDRY (2022): "Agricultural Technology in Africa," *Journal of Economic Perspectives*, 36, 33–56.

TAROZZI, A., A. MAHAJAN, B. BLACKBURN, D. KOPF, L. KRISHNAN, AND J. YOONG (2014): "Micro-Loans, Insecticide-Treated Bednets, and Malaria: Evidence from a Randomized Controlled Trial in Orissa, India," *American Economic Review*, 104, 1909–1941.

TINSLEY, E. AND N. AGAPITOVA (2018): "Private Sector Solutions to Helping Smallholders Succeed: Social Entreprise Business Models in the Agriculture Sector," Report, The World Bank.

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.

UDRY, C., F. DI BATTISTA, M. FOSU, M. GOLDSTEIN, A. GURBUZ, D. KARLAN, AND S. KOLAVALLI (2019): "Information, Market Access and Risk: Addressing Constraints to Agricultural Transformation in Northern Ghana," Draft Report.

URMINSKY, O., C. HANSEN, AND V. CHERNOZHUKOV (2016): "Using Double-Lasso Regression for Principled Variable Selection," 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.

WORLD BANK (2008): "World Development Report 2008: Agriculture for Development," Report, World Bank.

YOUNG, A. (2019): "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results*," *The Quarterly Journal of Economics*, 134, 557–598.

A Additional summary statistics and balance checks

A.1 Balance across various samples

Table A.1: Baseline balance by treatment assignment, primary sample

Variable	(1) Control Mean (SE)	(2) Treatment Mean (SE)	Difference (2)-(1)
Married (0/1)	0.90 (0.01)	0.89 (0.02)	-0.01
Household head has secondary school (0/1)	0.34 (0.02)	0.38 (0.03)	0.01
Household income >50% from farm labor (0/1)	0.79 (0.02)	0.76 (0.03)	-0.02
Used improved ag technology in 2016 (0/1)	0.62 (0.02)	0.66 (0.03)	0.00
Reports knowledge of 1AF practices (0/1)	0.06 (0.01)	0.14 (0.02)	0.07***
Intercropped maize and beans in 2016 (0/1)	0.47 (0.03)	0.56 (0.03)	0.05
Reports having credit access in 2016 (0/1)	0.71 (0.02)	0.73 (0.03)	0.03
Household size	6.63 (0.13)	6.75 (0.15)	0.20
Acres under maize cultivation in 2016	1.00 (0.04)	0.99 (0.05)	-0.01
Maize yield (kg/acre) in 2016	429.55 (21.81)	443.26 (21.75)	-26.34
F-statistic (test of joint significance)			1.25
Number of observations			682

Note: field office area fixed effects included in all balance test regressions. Standard errors are clustered at the treatment assignment (farmer group cluster) level.

Table A.2: Baseline balance by treatment assignment, pre-exposed sample

Variable	(1) Control Mean (SE)	(2) Treatment Mean (SE)	Difference (2)-(1)
Married (0/1)	0.86 (0.01)	0.88 (0.01)	0.01
Household head has secondary school (0/1)	0.41 (0.02)	0.47 (0.02)	0.06*
Household income >50% from farm labor (0/1)	0.77 (0.02)	0.79 (0.02)	0.01
Used improved ag technology in 2016 (0/1)	0.88 (0.01)	0.88 (0.01)	0.00
Reports knowledge of 1AF practices (0/1)	0.71 (0.02)	0.71 (0.02)	0.00
Intercropped maize and beans in 2016 (0/1)	0.48 (0.02)	0.43 (0.02)	-0.02
Reports having credit access in 2016 (0/1)	0.71 (0.02)	0.74 (0.02)	0.01
Household size	6.68 (0.10)	6.98 (0.11)	0.26*
Acres under maize cultivation in 2016	1.01 (0.03)	1.08 (0.04)	0.05
Maize yield (kg/acre) in 2016	602.12 (16.73)	661.56 (21.24)	79.04***
F-statistic (test of joint significance)			1.26
Number of observations			1214

Note: field office area fixed effects included in all balance test regressions. Standard errors are clustered at the treatment assignment (farmer group cluster) level.

A.2 Baseline comparison of pre-exposed and new farmers

Table A.3: Baseline balance across "primary" and pre-exposed samples

Variable	(1) Primary Sample Mean (SE)	(2) Pre-Exposed Sample Mean (SE)	Difference (2)-(1)
Panel A: Pre-Specified Control Variables			
Married (0/1)	0.90 (0.01)	0.87 (0.01)	-0.02
Household head has secondary school (0/1)	0.36 (0.02)	0.44 (0.01)	0.07**
Household income >50% from farm labor (0/1)	0.78 (0.02)	0.78 (0.01)	0.04*
Used improved ag technology in 2016 (0/1)	0.64 (0.02)	0.88 (0.01)	0.21***
Reports knowledge of 1AF practices (0/1)	0.10 (0.01)	0.71 (0.01)	0.59***
Intercropped maize and beans in 2016 (0/1)	0.51 (0.02)	0.46 (0.01)	-0.08***
Reports having credit access in 2016 (0/1)	0.72 (0.02)	0.73 (0.01)	0.06***
Household size	6.68 (0.10)	6.82 (0.07)	0.29**
Acres under maize cultivation in 2016	0.99 (0.03)	1.05 (0.02)	0.09**
Maize yield (kg/acre) in 2016	435.32 (15.59)	629.64 (13.34)	171.79***
Panel B: Additional Baseline Variables			
Female Respondent (0/1)	0.57 (0.02)	0.61 (0.01)	0.02
Income from non-farm labor (0/1)	0.50 (0.02)	0.49 (0.01)	-0.02
Income from business (0/1)	0.67 (0.02)	0.64 (0.01)	-0.01
Income from remittances (0/1)	0.43 (0.02)	0.51 (0.01)	0.09***
Income from formal employment (0/1)	0.26 (0.02)	0.26 (0.01)	0.01
Hired farm labor in 2016 (0/1)	0.72 (0.02)	0.75 (0.01)	0.04
DAP used in 2016 (kgs)	22.64 (1.45)	36.15 (1.03)	12.35***
# of non-maize crops cultivated	0.51 (0.03)	0.56 (0.02)	0.05
# extension officer visits, 2016	0.18 (0.04)	0.22 (0.03)	0.07
# of non-1AF farming org. memberships	0.17 (0.02)	0.15 (0.01)	0.02
Asset score	17.05 (0.27)	20.15 (0.20)	2.79***
F-statistic (test of joint significance)			39.19
Number of observations			1896

Note: field office area fixed effects included in all balance test regressions. Standard errors are clustered at the treatment assignment (farmer group cluster) level.

A.2.1 Distributions of nonbinary baseline variables by treatment status
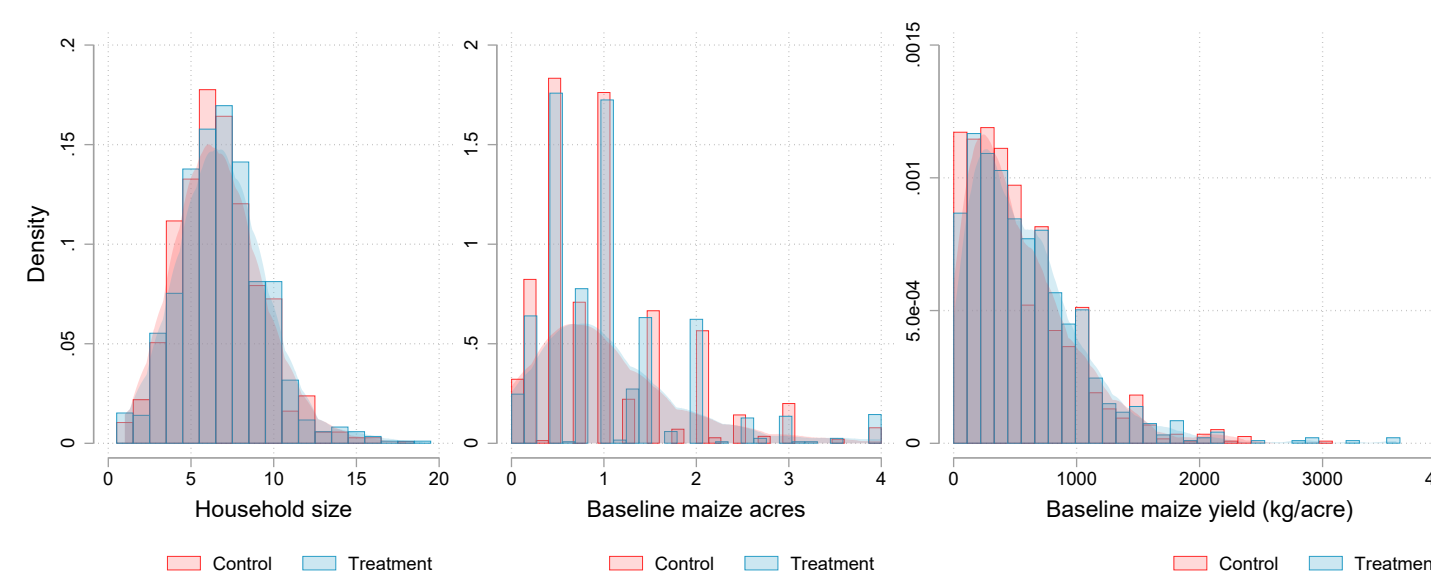

Figure A.1: Distributions of non-binary baseline characteristics, by treatment status

B   Variable construction and measurement

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.

Some readers may wish to see what the raw data look like, in addition to the regression results. Figure B.1 shows the distributions of productivity per acre and total maize output. In the top figure, for treatment farmers we show productivity on the portion of the land that farmers enrolled in the 1AF program. In the second figure, we show productivity on the remaining land which was not enrolled in the program. The bottom figure sums all maize produced by farmers to show the distribution of total maize output among farmers in each group.

We calculate projected revenues using average market prices from nearby vendors collected in September 2017. Following the PAP, we multiply prices by 1.08 to account for typical price increases over the consumption/selling season. This is a conservative assumption compared to findings in Burke, Bergquist, and Miguel (2019) who document maize price increases between 42% and 125% in post-harvest seasons from 2013 to 2017 in nearby areas of western Kenya. We calculate farmer costs using program costs and self-reported costs for treated farmers, and self-reported seed and fertilizer costs for control farmers. Labor costs include land prep, plowing, and planting costs, collected in a survey after planting, as well as post-planting costs collected at harvest time. For paid labor, we use farmer self-reported costs by planting phase. To include the opportunity cost of unpaid labor use, we calculate the mean day wage reported within the sample, devalue this mean wage by 50% (roughly the rural unemployment rate according to DHS data), and multiply this devalued mean by total person-days of unpaid labor for each planting phase. Profit is simply the difference between projected farmer revenues and costs.

Tables B.1 and B.2 measure fertilizer use in kilograms rather than valuing their use in USD. We detect a sizeable increase in fertilizer use when measured in kilograms, reinforcing that it is unlikely 1AF prices that are driving increased expenditures. We can also break down fertilizer use by phase, and here we see the underlying substitution behind the effect on fertilizer timing noted in Table 3. Farmers in the treatment group are not only using more fertilizer at the "correct" time, but also using less fertilizer at incorrect times.

Table B.1: Quantity of fertilizer used (kgs) by planting phase, primary sample

	At Planting		Post Planting	
	(1)	(2)	(3)	(4)
	DAP	CAN	DAP	CAN
1AF participant	22.250***	-0.054	-7.107***	15.410***
	(1.906)	(0.145)	(1.851)	(2.365)
Observations	682	682	682	682
R^2	0.236	0.013	0.220	0.211
Control Mean Dep. Var	8.976	0.213	10.256	17.980

Results in this table are from linear regressions of the outcomes in each column on the treatment dummy. Standard errors (in parentheses) clustered at the treatment assignment (farmer group cluster) level. All regressions include field office (site) fixed effects.

Table B.2: Quantity of fertilizer used (kgs) by planting phase, full sample

	At Planting		Post Planting	
	(1)	(2)	(3)	(4)
	DAP	CAN	DAP	CAN
1AF participant=1	22.201***	-0.035	-7.123***	15.724***
	(1.995)	(0.177)	(1.882)	(2.314)
Past 1AF participant=1	7.708***	0.235	-2.769*	5.527***
	(1.650)	(0.261)	(1.531)	(2.063)
1AF participant=1 × Past 1AF participant=1	-2.773	0.258	3.469*	0.406
	(2.555)	(0.326)	(1.811)	(3.003)
Observations	1896	1896	1896	1896
R^2	0.177	0.009	0.166	0.187
Control Mean Dep. Var	15.044	0.394	8.598	22.053

Results in this table are from linear regressions of the outcomes in each column on the treatment dummy. Standard errors (in parentheses) clustered at the treatment assignment (farmer group cluster) level. All regressions include field office (site) fixed effects.
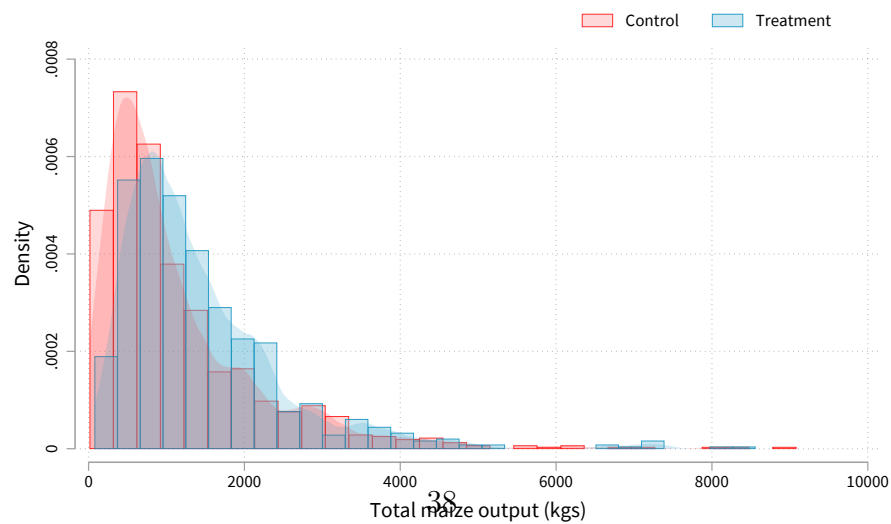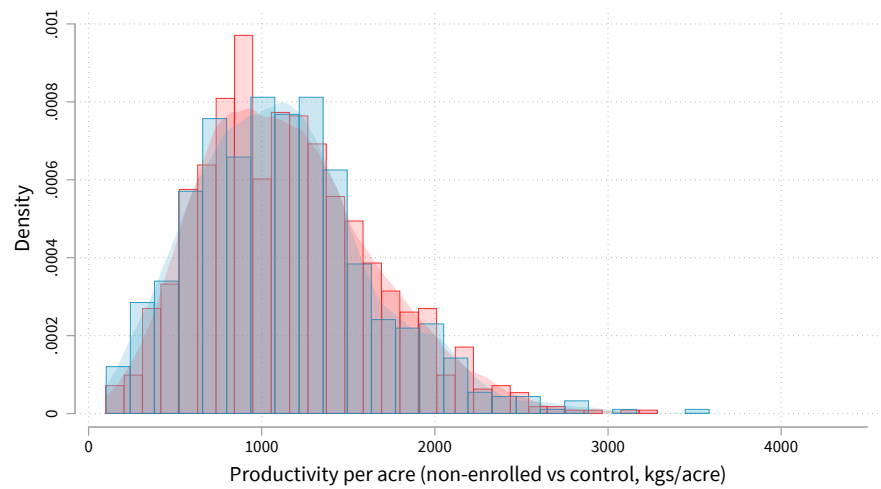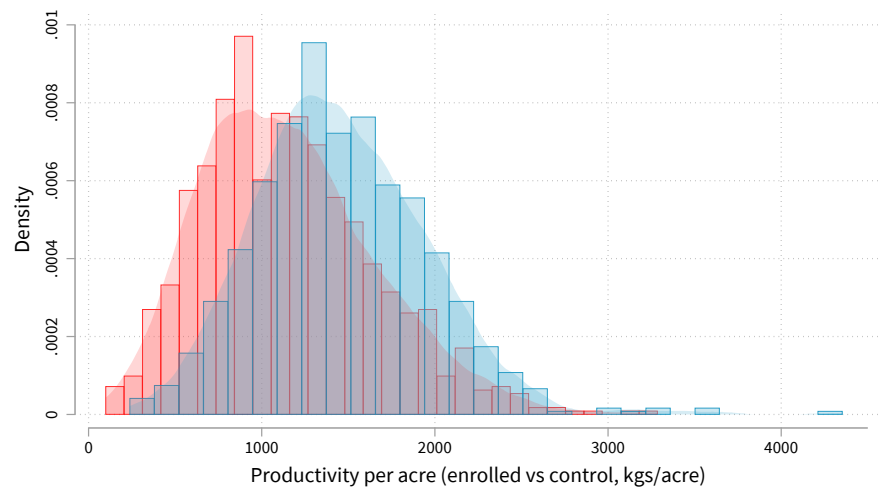
Figure B.1: Distributions of productivity and output

C Robustness and external validity

C.1 Attrition

This section addresses the different types of attrition that we find in our data: first, we examine attrition that occurred between the baseline survey but before farmers passed the pre-enrollment qualification. To qualify for the study, farmers had to pay a 500 KES deposit and form farmer groups. Farmers who did not complete the prepayment and/or failed to form a group of at least 3 members were dropped. Because the baseline was completed before this qualification stage, the study intentionally sampled more farmers than the desired sample size. Since treatment was assigned after the pre-qualification, we should not expect any threats to internal validity from post-baseline attrition. Understanding what caused this attrition may however be informative with respect to generalizability. We examine this in Section C.1.1.

Conditional on being enrolled in the study, attrition is driven by two main factors: missing land size values and missing harvest data. During data cleaning, we lose 175 observations that are missing land size and 329 observations that are missing dry weight harvest data (of these 329, we observe fresh weight harvest measurements for 53 observations). This attrition is not trivial, and in the following subsections we discuss how we ensure our results are robust to a variety of assumptions about the missing data.

C.1.1 Attrition prior to season and external validity

Table C.1 shows how the qualified and dropped samples differ. The farmers who managed to qualify differ significantly from the farmers who dropped out. Farmers who passed the pre-qualification stage were more likely to be pre-exposed, had better knowledge of 1AF practices, were more likely to use improved seeds and fertilizer, and were more likely to intercrop. Qualified farmers also seem more specialized in farming, as measured by the higher likelihood of receiving more than half their income from farm labor, farming more acres for maize in 2016 and 2015, and the amount of maize that they harvested in previous seasons. Finally, qualified farmers are wealthier, as indicated by ownership of more land and a higher asset score.

These variables suggest that 1AF may not reach the poorest farmers—a common challenge for entrepreneurially-focused agriculture programs (Carter, Tjernström, and Toledo, 2019). That said, compared to a more representative sample—Tegemeo Institute's panel survey of maize farming households—the qualified farmers cultivate fewer acres than the

Table C.1: Baseline balance across enrolled and dropped groups

Variable	(1) Qualified Mean (SE)	(2) Dropped Mean (SE)	Difference (2)-(1)
Panel A: Pre-Specified Control Variables			
Married (0/1)	0.88 (0.01)	0.88 (0.01)	0.00
Household head has secondary school (0/1)	0.40 (0.01)	0.32 (0.02)	-0.08***
Household income >50% from farm labor (0/1)	0.77 (0.01)	0.72 (0.02)	-0.05**
Used improved ag technology in 2016 (0/1)	0.77 (0.01)	0.68 (0.02)	-0.09***
Reports knowledge of 1AF practices (0/1)	0.49 (0.01)	0.34 (0.02)	-0.15***
Intercropped maize and beans in 2016 (0/1)	0.47 (0.01)	0.43 (0.02)	-0.04*
Reports having credit access in 2016 (0/1)	0.72 (0.01)	0.70 (0.02)	-0.02
Household size	6.72 (0.05)	6.09 (0.10)	-0.63***
Acres under maize cultivation in 2016	1.01 (0.02)	0.84 (0.03)	-0.17***
Maize yield (kg/acre) in 2016	555.75 (9.17)	470.22 (17.01)	-85.52***
Panel B: Additional Baseline Variables			
Female Respondent (0/1)	0.59 (0.01)	0.51 (0.02)	-0.08***
Income from non-farm labor (0/1)	0.48 (0.01)	0.57 (0.02)	0.10***
Income from business (0/1)	0.64 (0.01)	0.58 (0.02)	-0.06***
Income from remittances (0/1)	0.48 (0.01)	0.46 (0.02)	-0.02
Income from formal employment (0/1)	0.26 (0.01)	0.19 (0.02)	-0.08***
Hired farm labor in 2016 (0/1)	0.72 (0.01)	0.67 (0.02)	-0.05**
DAP used in 2016 (kgs)	30.83 (0.75)	22.69 (1.92)	-8.14***
# of non-maize crops cultivated	0.54 (0.01)	0.46 (0.03)	-0.07**
# extension officer visits, 2016	0.19 (0.02)	0.13 (0.03)	-0.06
# of non-1AF farming org. memberships	0.15 (0.01)	0.10 (0.01)	-0.05**
Asset score	18.75 (0.15)	16.18 (0.26)	-2.56***
F-statistic (test of joint significance)			94.12
Number of observations			3002

Tegemeo sample (see Figure C.1).[22] Note that the Tegemeo survey relies on self-reported acreage; we therefore report our baseline self-reported acreage variable (dotted line in Figure C.1) in addition to the GPS-measured land sized (short-dashed line in Figure C.1). Our sample farmers farm smaller plots than the average TAPRA survey participant.

Our pre-exposed farmers could be driving these differences. To check this, Table C.2 repeats the balance test for the "primary" sample (i.e., without the pre-exposed farmers). Farmers who make it past the pre-qualification stage are still significantly different, but mostly related to wealth and agricultural specialization. In terms of knowledge of 1AF prac-

---

[22]Tegemeo Institute's panel survey was designed to be broadly representative of the maize-growing regions of Kenya.

Table C.2: Baseline balance across enrolled and dropped groups among primary sample

Variable	(1) Qualified Mean (SE)	(2) Dropped Mean (SE)	Difference (2)-(1)
Panel A: Pre-Specified Control Variables			
Married (0/1)	0.88 (0.01)	0.87 (0.02)	-0.01
Household head has secondary school (0/1)	0.35 (0.02)	0.27 (0.03)	-0.08***
Household income >50% from farm labor (0/1)	0.76 (0.01)	0.68 (0.03)	-0.08***
Used improved ag technology in 2016 (0/1)	0.59 (0.02)	0.57 (0.03)	-0.02
Reports knowledge of 1AF practices (0/1)	0.10 (0.01)	0.10 (0.02)	0.01
Intercropped maize and beans in 2016 (0/1)	0.51 (0.02)	0.47 (0.03)	-0.04
Reports having credit access in 2016 (0/1)	0.71 (0.02)	0.68 (0.03)	-0.02
Household size	6.59 (0.09)	5.84 (0.13)	-0.76***
Acres under maize cultivation in 2016	0.95 (0.03)	0.76 (0.04)	-0.19***
Maize yield (kg/acre) in 2016	433.24 (13.78)	398.07 (22.38)	-35.17
Panel B: Additional Baseline Variables			
Female Respondent (0/1)	0.57 (0.02)	0.52 (0.03)	-0.05
Income from non-farm labor (0/1)	0.46 (0.02)	0.56 (0.03)	0.10***
Income from business (0/1)	0.63 (0.02)	0.58 (0.03)	-0.05*
Income from remittances (0/1)	0.42 (0.02)	0.47 (0.03)	0.05
Income from formal employment (0/1)	0.25 (0.01)	0.18 (0.02)	-0.07***
Hired farm labor in 2016 (0/1)	0.68 (0.02)	0.67 (0.03)	-0.01
DAP used in 2016 (kgs)	21.92 (1.28)	14.82 (1.75)	-7.10***
# of non-maize crops cultivated	0.51 (0.02)	0.44 (0.04)	-0.06
# extension officer visits, 2016	0.16 (0.03)	0.14 (0.05)	-0.02
# of non-1AF farming org. memberships	0.15 (0.02)	0.09 (0.02)	-0.06**
Asset score	16.31 (0.26)	14.79 (0.35)	-1.51***
F-statistic (test of joint significance)			46.52
Number of observations			1155

tices, use of improved seed and fertilizer, credit access, and intercropping, the two groups are statistically indistinguishable. We see also that many differences, while significant, are much smaller in magnitude.
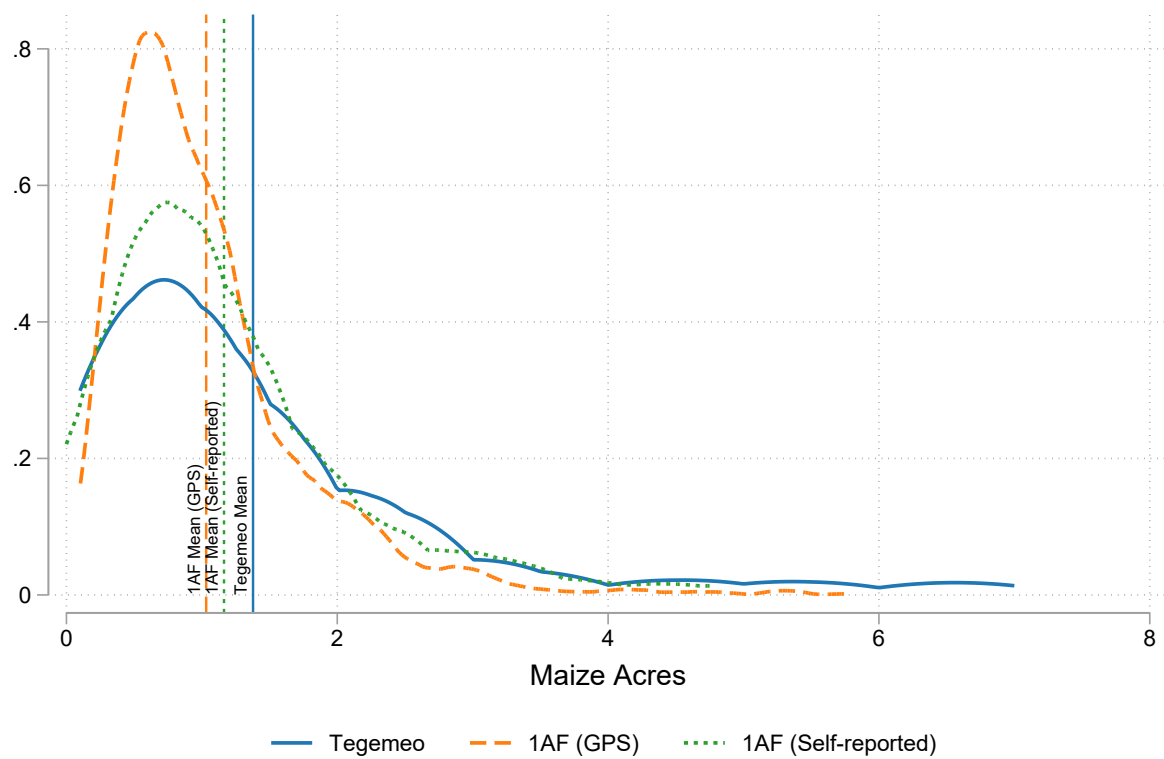
Figure C.1: Comparison of maize acres across study sample and TAPRA survey

C.1.2   Within-season attrition

To test the robustness of our results to the potentially non-random missing data, we impute missing values in a number of ways. The main source of missingness occurs in the harvest variable (yields on the crop-cut plot) and the land size variable.

For yields, we use two different imputation strategies. First, for a subset of farmers, we have non-missing "fresh weight" yields (but missing dry weight yields). We use the sample of farmers for whom we have both fresh and dry maize weights to estimate the relationship between the two variables. We then predict the dry weight yields for the subsample with only fresh weights. For farmers without any harvest weights, we make the most conservative assumption and assume zero yields (i.e., total crop failure). This should allow us to estimate a realistic lower bound on possible treatment effects. Second, we complement the above with a simpler approach where we replace missing yield values with the mean yields by plot type (treated-enrolled plot, treated-non-enrolled plot, control plot).

For missing land size, we use survey-collected self-reported measures across three different surveys that were part of the data collection efforts. Similarly to the yield imputations, we use the estimated relationships in the sample of households with overlapping GPS-measured land size and self-reported data to predict land size for the missing observations. Here, too, we conduct the predictions separately by plot type. Given that we have several different self-reported measures, we report results from four different imputations, using each self-reported value on its own *(i)-(iii)*, and using the mean of these three self-reports *(iv)*. For the latter, if a participant is missing one of the three self-reports, we take the mean of the non-missing self-reported values.

Figures C.2 and C.3 show the results of testing these various strategies for dealing with attrition in our two key variables.

Figure C.2: Treatment effect estimates on maize productivity per acre with various methods of accounting for within-season attrition
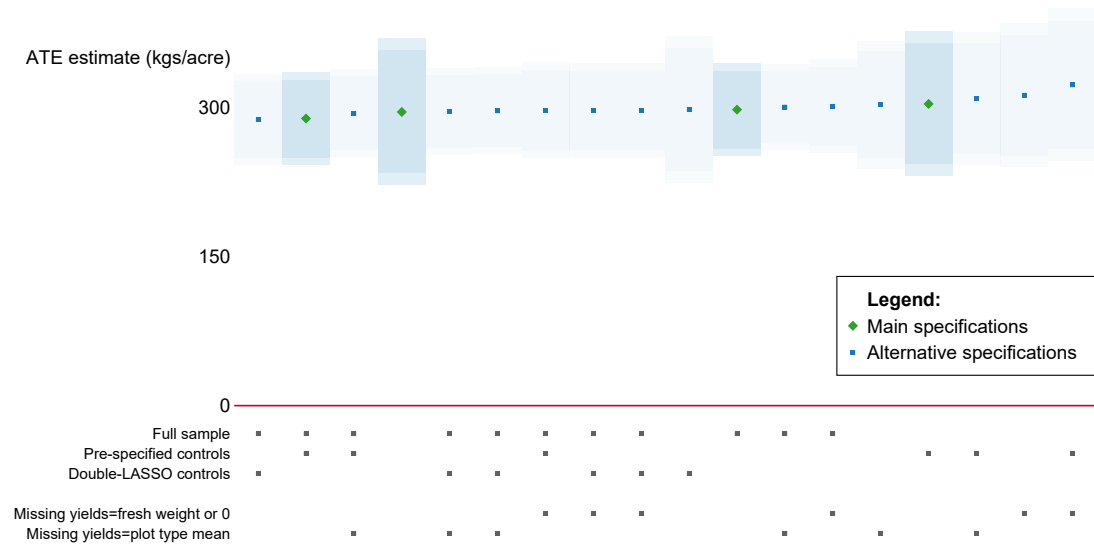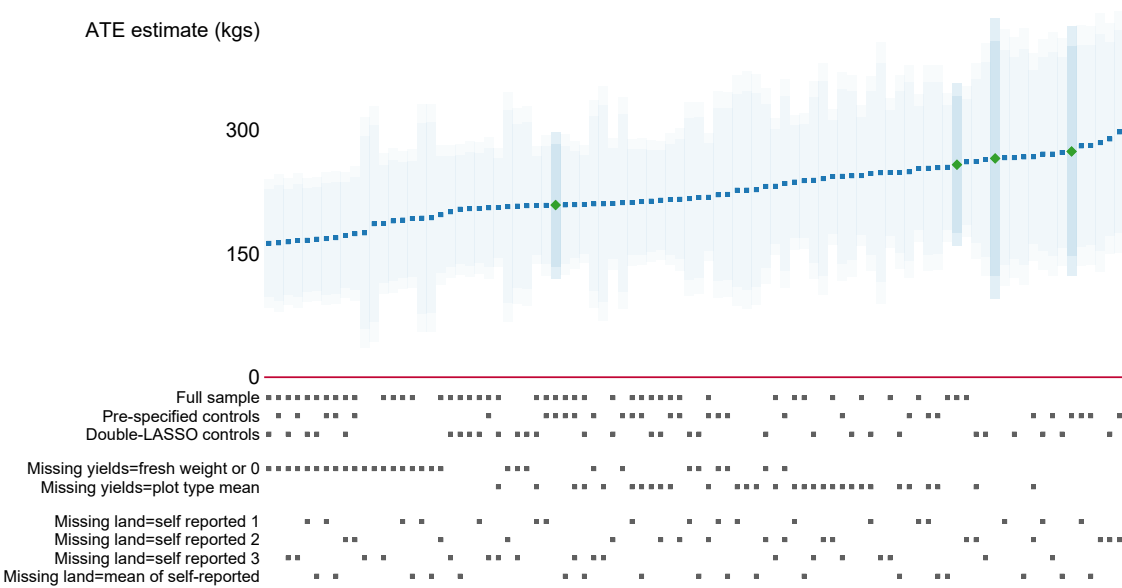

Figure C.3: Treatment effect estimates on total maize output with various methods of accounting for within-season attrition

C.2 Alternative specification details

Table C.3: Productivity and output with pre-specified controls

	Productivity per acre enrolled vs. control		Productivity per acre non-enrolled vs. control		Total output per farmer	
	(1)	(2)	(3)	(4)	(5)	(6)
1AF participant	303.99***	308.68***	-51.18	-42.24	274.15***	270.66***
	(36.810)	(36.910)	(35.680)	(39.250)	(75.820)	(81.870)
Pre-exposed		-9.72		-29.31		70.48
		(37.800)		(37.510)		(71.780)
1AF participant		-29.67		-20.64		-93.74
× pre-exposed		(49.940)		(54.990)		(113.330)
1AF + (1AF × Pre-exposed)		279.01***		-62.88		176.92***
		(31.31)		(41.25)		(62.34)
Control group mean	1128.39	1150.49	1128.39	1150.49	1082.38	1164.00
Observations	682	1896	614	1701	682	1896

Note: This table presents results from linear regressions of the outcomes at the top of each column on the treatment dummy. Productivity per acre is measured as kgs of maize per acre cultivated, and total maize output is measured in kgs. In columns 1 and 2, enrolled refers to the land farmers in the treatment group enrolled in the 1AF program, which may be less than the full amount of land on which they cultivate maize. In columns 3 and 4, we include only farmers who also cultivated maize on land they did not enroll with 1AF and compare the yields from that non-enrolled land to yields from control group farmers. Standard errors (in parentheses) are clustered at the treatment assignment (farmer group cluster) level. Columns 1, 3, and 5 include only farmers who had never previously participated in the 1AF program. Columns 2, 4, and 6 additionally include a sample of farmers who had previously enrolled in the 1AF program (see Section 2.2). All regressions include strata (field office) fixed effects. This table replicates Table 2 above but includes the set of pre-specified controls. Binary baseline controls are married, at least secondary education by male household head, whether the household receives half of its income from farm labor, whether the household used improved seeds at baseline, whether the household reported knowledge of 1AF practices at baseline, whether the farmer intercropped beans and maize at baseline, and whether the household reported having access to credit. Continuous baseline controls are household size, a measure of Fall Army Worm presence collected during the 2017 season, baseline maize acres, and baseline maize yields per acre.

Although the PAP pre-specified a set of baseline controls that should be included in analysis of our outcomes of interest, the baseline survey contains a number of additional covariates that may be useful in improving the precision of treatment effect estimates.[23] Following Belloni et al. (2014) and Urminsky et al. (2016), we implement a double-Lasso

---

[23]An additional deviation from the original PAP is that the PAP included an additional covariate measuring Fall Army Worm incidence. During the course of the 2017 season, a pest called Fall Army Worm had a dramatic effect on Kenyan farmers. We were concerned that the pest may affect treatment estimates. To measure the extent of FAW, enumerators visited farmer fields during the growing season and randomly selected 30 plants to inspect and check for signs of Fall Army Worm. The problem was indeed widespread: nearly 80% of farmers had at least some plants affected by FAW, and 66% of farmers had signs of FAW on all inspected plants. In theory, we may be concerned that FAW incidence could be higher among treated farmers, for example if their maize is healthier and more easily allows FAW to propagate. In practice, however, including this variable as an additional control has no meaningful impact on the magnitude or precision of our treatment effect estimates.

Table C.4: Profits with pre-specified controls

	Profit (PAP Definition)		Profit (Full Labor Costs)	
	(1)	(2)	(3)	(4)
1AF participant	61.93**	59.08*	56.75**	54.09*
	(27.520)	(30.230)	(27.430)	(30.040)
Pre-exposed		28.79		29.72
		(26.640)		(26.340)
1AF participant		-39.38		-40.66
× pre-exposed		(41.180)		(40.900)
1AF + (1AF × Pre-exposed)		19.70		13.42
		(22.98)		(22.84)
Control group mean	335.35	364.61	320.70	350.92
Observations	682	1896	682	1896

Note: This table presents results from linear regressions of the outcomes shown at the top of each column on the treatment dummy. Profits are calculated as the value of maize production less costs (see Appendix B), and are reported in USD. Columns 1 and 2 follow the PAP and account for unpaid labor costs by valuing unpaid time at 50% of the average day-wage reported in the sample for paid labor. Columns 3 and 4 instead value unpaid labor at the full average day-wage reported in the sample. Standard errors (in parentheses) are clustered at the treatment assignment (farmer group cluster) level. Odd columns include only farmers who had never previously participated in the 1AF program. Even columns additionally include a sample of farmers who had previously enrolled in the 1AF program (see Section 2.2). All regressions include strata (field office) fixed effects. This table replicates Table 6 above but includes the set of pre-specified controls. Binary baseline controls are married, at least secondary education by male household head, whether the household receives half of its income from farm labor, whether the household used improved seeds at baseline, whether the household reported knowledge of 1AF practices at baseline, whether the farmer intercropped beans and maize at baseline, and whether the household reported having access to credit. Continuous baseline controls are household size, a measure of Fall Army Worm presence collected during the 2017 season, baseline maize acres, and baseline maize yields per acre.

procedure for variable selection. In addition to all pre-specified controls and fixed effects, we additionally include respondent gender, DAP use, crop diversity, extension officer visits, membership in other farming organizations, use of hired farm labor, an assets score, and dummies for income from nonfarm labor, businesses, remittances, or formal employment. Variable selection is distinct for each outcome variable and each sample definition. The results of this exercise, presented in Table C.5, demonstrate that the results are not particularly sensitive to the inclusion of an optimal set of controls.

Table C.5: Productivity and output with double-Lasso-selected controls

	Productivity per acre enrolled vs. control		Productivity per acre non-enrolled vs. control		Total output per farmer	
	(1)	(2)	(3)	(4)	(5)	(6)
1AF participant	299.67***	306.46***	-53.32	-37.69	266.08***	253.09***
	(37.190)	(36.580)	(38.080)	(40.260)	(73.720)	(77.590)
Pre-exposed		-7.57		-20.40		62.58
		(37.820)		(37.430)		(69.960)
1AF participant		-24.46		-24.68		-72.86
× pre-exposed		(49.660)		(56.590)		(109.440)
1AF + (1AF × Pre-exposed)		282.00***		-62.38		180.23***
		(31.15)		(40.73)		(59.68)
Control group mean	1128.39	1150.49	1128.39	1150.49	1082.38	1164.00
Observations	682	1896	614	1701	682	1896

Note: This table presents results from linear regressions of the outcomes at the top of each column on the treatment dummy. Productivity per acre is measured as kgs of maize per acre cultivated, and total maize output is measured in kgs. In columns 1 and 2, enrolled refers to the land farmers in the treatment group enrolled in the 1AF program, which may be less than the full amount of land on which they cultivate maize. In columns 3 and 4, we include only farmers who also cultivated maize on land they did not enroll with 1AF and compare the yields from that non-enrolled land to yields from control group farmers. Standard errors (in parentheses) are clustered at the treatment assignment (farmer group cluster) level. Columns 1, 3, and 5 include only farmers who had never previously participated in the 1AF program. Columns 2, 4, and 6 additionally include a sample of farmers who had previously enrolled in the 1AF program (see Section 2.2). All regressions include strata (field office) fixed effects. This table replicates Table 2 above but includes controls selected with double-Lasso-selected controls (Belloni et al., 2014). The Lasso regressions select from all baseline covariates included in Table C.3, as well as the following additional baseline covariates: respondent gender, DAP use, crop diversity, extension officer visits, membership in other farming organizations, use of hired farm labor, an assets score, and dummies for income from nonfarm labor, businesses, remittances, or formal employment. Variable selection is distinct for each outcome variable and each sample definition.

Table C.6: Profits with double-Lasso-selected controls

	Profit (PAP Definition)		Profit (Full Labor Costs)	
	(1)	(2)	(3)	(4)
1AF participant	59.69**	53.02*	54.32**	47.84*
	(27.070)	(28.970)	(26.940)	(28.760)
Pre-exposed		25.70		26.58
		(26.110)		(25.820)
1AF participant		-31.33		-32.44
× pre-exposed		(39.950)		(39.650)
1AF + (1AF × Pre-exposed)		21.68		15.41
		(22.16)		(22.03)
Control group mean	335.35	364.61	320.70	350.92
Observations	682	1896	682	1896

Note: This table presents results from linear regressions of the outcomes shown at the top of each column on the treatment dummy. Profits are calculated as the value of maize production less costs (see Appendix B), and are reported in USD. Columns 1 and 2 follow the PAP and account for unpaid labor costs by valuing unpaid time at 50% of the average day-wage reported in the sample for paid labor. Columns 3 and 4 instead value unpaid labor at the full average day-wage reported in the sample. Standard errors (in parentheses) are clustered at the treatment assignment (farmer group cluster) level. Odd columns include only farmers who had never previously participated in the 1AF program. Even columns additionally include a sample of farmers who had previously enrolled in the 1AF program (see Section 2.2). All regressions include strata (field office) fixed effects. The Lasso regressions select from all baseline covariates included in Table C.4, as well as the following additional baseline covariates: respondent gender, DAP use, crop diversity, extension officer visits, membership in other farming organizations, use of hired farm labor, an assets score, and dummies for income from nonfarm labor, businesses, remittances, or formal employment. Variable selection is distinct for each outcome variable and each sample definition.

D   Additional heterogeneity results

Beyond the heterogeneity results found in Section 5.3, another way to view heterogeneity is through distribution regressions (Chernozhukov, Fernández-Val, and Melly, 2013). These are related to unconditional quantile partial effects but can be easier to interpret since the x-axis shows meaningful values rather than percentiles of a distribution. We implement this by running a series of regressions of the following form

$$1(Y > x)_{ij} = \alpha + \beta_x T_i + \gamma_j + \epsilon_i,$$

where Y is the outcome of interest, x varies along the x-axis of each figure along the support of the outcome variable, T_i is the treatment dummy, and γ_j is a cluster fixed effect. Each blue dot in Figure D.1 is an estimated β_x coefficient. The results shown here are from estimations with a linear probability model, but the results are robust to using a logit model to estimate the threshold probabilities.

The results for maize productivity per acre and total output both suggest that the effects are largest around the lower end of the distribution and they attenuate at large values of the outcomes. Maize yields likely have some physiological upper bound, beyond which decreasing marginal returns to inputs start to make additional intensification less effective absent further investment in mechanization or land consolidation.

A key difference between the Chernozhukov et al. (2018) approach and the distribution regression approach of Chernozhukov et al. (2013) is that the former focuses on whether specific covariates can predict the size of participants' treatment effects. Figure 1 presents the Sorted Group Average Treatment Effects (GATES) estimated using Neural Nets. Each vertical bar represents the estimated treatment effect at a different percentile of the predicted treatment effect distribution.

For the groups who are predicted to have low treatment effects based on observables, the GATES estimate is only significantly greater than zero for program maize yields. That said, the least-affected and most-affected groups do not differ starkly from each other and the method introduces substantial noise in our relatively small sample. While we do not detect much heterogeneity along the distribution of the outcome variable, this does not automatically rule out the existence of subgroups for whom the treatment is more or less effective. Since this approach relies on covariates to predict heterogeneity, it could nevertheless be the case that we are not including the correct covariates. We test the method with both our pre-specified set of controls and a larger set of covariates which we also use for our double-Lasso methods discussed above in Appendix C.2. In both cases we

fail to reject the null hypothesis that the heterogeneity loading parameter differs from zero.
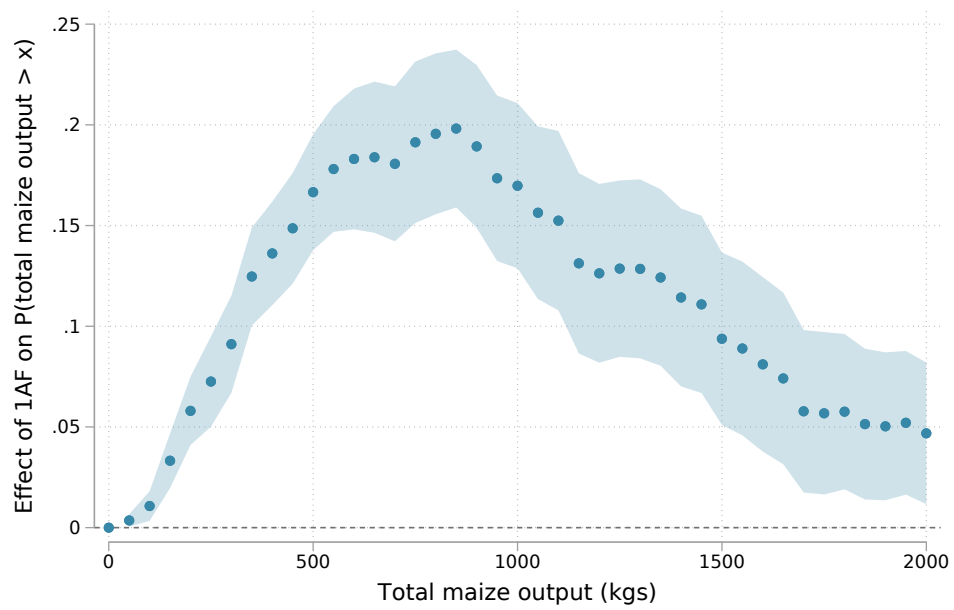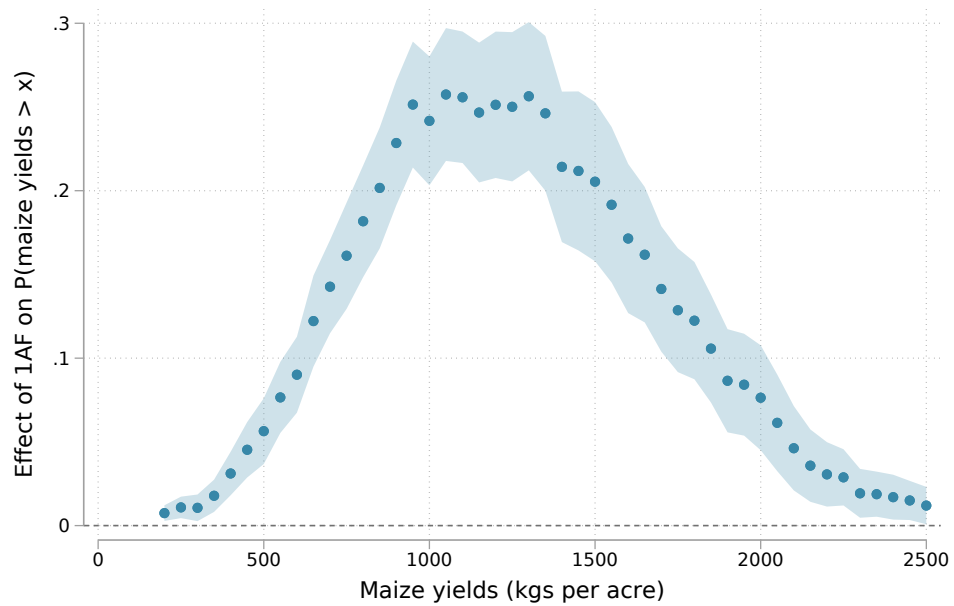
Figure D.1: Distribution regression results

E   Audit procedures

To boost the credibility of the data collection, 1AF contracted Intermedia Development Consultants (iDC), an independent survey firm, to carry out a three-step audit of the data collection. The full iDC audit report is available from the authors upon request; the overall conclusion of the audit stated the following "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."

The goal of the audit exercise was to evaluate 1AF against best-practice M&E standards, on dimensions including the technical capacity of monitoring staff, supervision, the M&E strategy and planning, data collection, data quality (accuracy, reliability and validity), data recording and analysis, dissemination and use of data, and participation of stakeholders. The auditors recorded observations in the field, which were then compared with data collected by the enumerators to expose any gaps or strengths in the OAF system.

The audit procedure took place in three steps: Step 1 reviewed the planting compliance and crop mix survey data collection. The audit covered a sample of the 2,425 farmers surveyed. Step 2 covered data collection of the beans harvest survey, covering a sample of around 1,400 farmers. Step 3 entailed completing similar activities for the maize harvest survey, covering a sample of around 2,425 farmers.

As part of this process, the audit team participated in group meetings held by the 1AF monitoring team and 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. The audit team observed the work of supervisors on 11 separate 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.

Table E.1: Materials reviewed during iDC audit

Step 1	Step 2	Step 3
Back-check code book	Questionnaires: audited beans dry weight survey	Questionnaire: Harvest survey back-check
Planting compliance and crop mix code book	Questionnaires: for audited bean fresh weight survey	Maize box check
Back-check strategy	Questionnaires: audited harvest box survey	Fresh maize harvest questionnaire
Data collection process and flow	Teso Trial baseline report	Dry weight harvest questionnaire
Teso Trial sites	Teso Trial design and analysis plan	Supervisors checklist
Updated guidance on back-checks	feedback Receipt for beans harvest	List of enumerators and their contacts
Kenya feedback receipts		List of back-checkers and their contacts
Teso Trial baseline report		List of supervisors and contacts
Analysis and design plan		Teso Trial planting compliance crop mix back-check analysis
Audited cases April 4-15		