One Acre Fund Tree Program: Midline evaluation report

Key findings

SEPTEMBER 2020

laterite

DATA | RESEARCH | ADVISORY

One Acre Fund Tree Program: Midline evaluation report. Key findings

September 2020

www.laterite.com

Please cite this report as:

Laterite. (2020). One Acre Fund Tree Program: Midline evaluation report. Nairobi, Kenya.

ACKNOWLEDGEMENTS

Laterite would like to thank One Acre Fund for funding this project and especially Kim Siegal and Haley Kawaja, for the time and dedication to this project.

Laterite also would like to sincerely thank the One Acre Fund clients that participated in the interviews for their insights.

Laterite would like to acknowledge our field team for their time, assistance and insights during this project. We would not have completed this assignment without their hard work and perseverance.

Finally, we also acknowledge the support of Laterite colleagues Tessa Ahner-McHaffie and Anne Wagaturi as well as our field supervisors Winnie Kimiti and Martin Gichuru.

This report was delivered by Laterite's team composed of Sylvia Onchaga, Francisco Garcia, and Dimitri Stoelinga.

Table of contents

List of abbreviations

1AF	One Acre Fund
(L)ATE	(Local) Average Treatment Effect
ITT	Intention to Treat
IV	Instrumental Variable
ITT	Intention to Treat
Ksh.	Kenyan Shillings
(C)RCT	(Cluster) Randomized Control Trial
TLUs	Tropical Livestock Units

List of figures

List of tables

1 Introduction

One Acre Fund (1AF) has been supporting smallholder farmers in Kenya through the provision of agricultural training and financing since 2006. The financing component involves providing farmers with farm inputs on credit at the time of planting, who then make repayments to 1AF in installments over the season. The inputs primarily constitute a core package that contains improved maize seeds, bean seeds, fertilizer and a range of add-on products such as solar lights, vegetable seeds, improved crop storage bags, cook stoves, and sanitary pads.

In 2010, 1AF introduced a Tree Program to their core loan package with the aim of increasing the client farmers' tree assets as a source of extra income from their sale. The Tree Program contains grevillea seeds, planting fertilizer, large planting the seeds, smaller planting bags for raising individual seedlings once they are big enough, and a set of trainings specifically on tree planting and maintenance.

Laterite was contracted by 1AF in 2018 to undertake a randomized control trial (RCT) to evaluate the impact of the Tree Program in the Kericho and Uasin Gishu counties[1] of Kenya, where 1AF expanded their program in the 2019 long rains season. The expansion of 1AF operations into new areas of Kenya gave a unique opportunity for an impact evaluation that is not compromised to pre-exposure to the program. Laterite is responsible for all stages of the impact evaluation, from design through data collection, data cleaning, analysis and reporting.

This report addresses the midline stage of the impact evaluation. The midline data was collected between the 3rd of February 2020 and the 6th March 2020 from the farmers interviewed at baseline that consented to participate in the study. The purpose of the midline survey was to assess changes in the uptake of tree-planting activities, survival rates of planted grevillea trees, perceptions, and attitudes of farmers towards tree-planting, and farmers' knowledge of tree planting best practices. In total, 1,730 farmers out of 1,852 farmers interviewed at baseline responded to the midline survey, representing a 93% response rate[2].

This report provides the key midline evaluation findings which are organized to align with the research objectives and hypotheses. We note that there are limitations to the internal and external validity this study. Among risks to internal validity identified at the pre-analysis stage, we encounter an attrition rate of 7%, a non-compliance rate of 8% for treatment farmers, and a 3% rate for control farmers receiving tree kits at midline.

We also encounter limitations to the external validity, outlined by One Acre Fund. First, 1AF finds in their internal reports that the 2019 expansion areas are not fully similar to other

[1] For the purposes of the 1AF program, the expansion area comprises Kipkelion- and Kabiyet districts. Kenya's administrative divisions were adjusted under the 2010 constitution.
[2] The Laterite team had organized a mop-up data collection exercise in mid-March to attempt to interview baseline farmers that could not be reached during the main data collection, but this was cancelled due to the restrictions imposed by the Government of Kenya as a response to the COVID-19 pandemic.

program areas. In general, farmers in the Kericho and Uasin Gishu expansion areas grow more trees and there are a number of factories in Kericho that buy eucalyptus trees as fuel. The higher number of pre-existing timber trees might limit our ability to assess an incremental impact on the number of timber trees planted. Second, in 2019 the input delivery was delayed due to supply issues, which led to alterations to the tree training schedule. 1AF suspects that some training was rushed or not delivered at all. Third, the program was affected by seasonal effects. The long rains came late in 2019, with a “false start” to the rainy season. Around a quarter of 1AF farmers had to replace their field with new seed because of this. Finally, the 1AF Tree Program has changed since this study began. 1AF is now distributing eucalyptus trees in areas that are safe to plant and the organization is trialing distributing seedlings rather than seeds. These limitations constrain our ability to generalize the findings of this report to the overall Tree Program.

2 Research Objectives

Over the course of this three-year study we aim to look at the impact that the 1AF Tree Program has in terms of the uptake of tree planting activities, the survival rates of planted trees, the financial value of tree assets, and the perceptions and attitudes of farmers towards tree planting. The specific research questions are listed below.

2.1 Primary Research Questions

The primary research questions to be addressed include:

1. **Uptake and survival:** Do treatment farmers have additional trees after the two years compared to control farmers; how many additional trees do they have, on average; and, what does this translate to in terms of current and future potential financial value?
2. **Cost-Benefit:** What is the overall value to the treatment farmer per tree planted considering opportunity costs regarding land-use, time-use and cash-use? Is there a net change (increase or decrease) in the value of total tradeable assets, combining trees and livestock, for the treatment farmers over the course of the two years compared to control farmers?
3. **Attitudes:** Is there a change in farmers' perceptions and attitudes towards tree planting activities and farmers' understanding of best practices around tree-planting that can be attributed to the 1AF Tree Program?

2.1 Secondary Research Questions

Secondary research questions include:

1. What are the socio-economic determinants of tree-planting?
2. To what extent do tree survival rates among treatment farmers correlate with the level of adoption/application of best practices for tree-planting outlined in the training sessions?
3. Does a farmer's knowledge of best practices for tree-planting and tree-maintenance improve as a result of the tree-program?
4. Does planting trees one year preclude planting trees in subsequent years? Does receiving a tree-kit and training in 2019 correlate positively or negatively with tree planting in 2020?
5. Does the pattern of land allocation to different resources on the farm change as a result of the tree-program?
6. What are the usage patterns of trees among our treatment and control groups? Are there any changes over time? What are the main differences between the different tree types?

7. Does the tree kit intervention affect household's time/labor allocation across farming and household chores?
8. Does a farmer's perception of a tree's financial value change as a result of the Tree Program? How does this compare to valuations by tree-traders and control farmers not exposed to 1AF tree value messaging?

3 Study Methodology

3.1 Evaluation Design and Sample

The research design is based on a cluster-randomized controlled trial. The unit of randomization is the training group that a farmer is assigned to and the unit of measurement is individual farmers.

At baseline, the sampling frame was all 1AF farmers in the new expansion sites, where 1AF started operating in 2019. Based on power calculations to identify the minimum detectable effect on key outcomes of interest, 226 training groups in 37 1AF training sites were assigned to the treatment and control arms. Overall, 1,852 farmers from these groups were sampled and interviewed at baseline.

The sampling frame for the midline stage of this study comprised all the households where a farmer had been interviewed at baseline. Farmers within the same household different than the respondent from the baseline survey were interviewed in cases where the baseline farmer could not be interviewed throughout the data collection period.

3.1.1 1AF Intervention

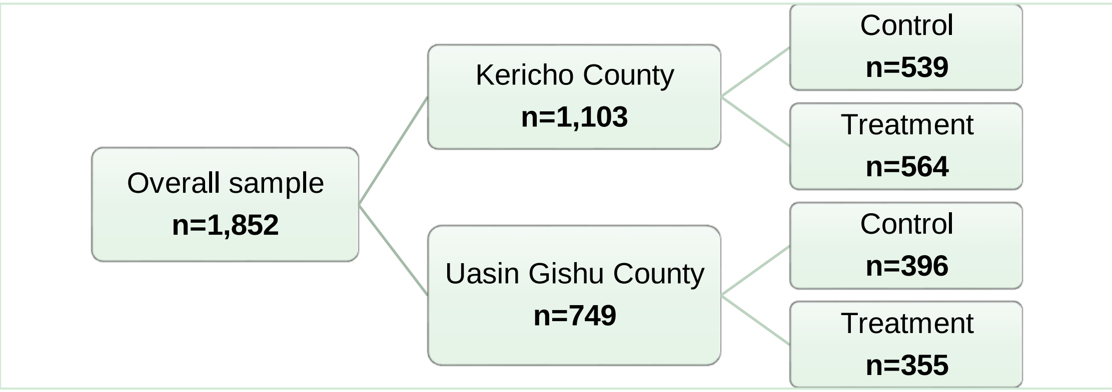
Farmers in the treatment training groups were scheduled to receive the "Tree Program". This consists of tree kits, and tree training as part of their regular base package of inputs (improved seeds and fertilizer for maize and beans plus any add-on products) for the 2019 long rains season. Control training groups were scheduled to receive a base package that did not include the tree kits and tree-specific training.

The underlying assumption for this evaluation is that treatment and control farmers differ from each other only by chance and that differences in outcomes of interest are explained by the treatment. At baseline, we tested this hypothesis and concluded that allocation to treatment is not correlated with outcome variables of interest or demographic characteristics.

3.1.2 Sample Size

A sample of 1,852 farmers (Figure 1) was determined to be sufficient to detect at least a 33.0% average difference in tree assets between the treatment and control groups at 0.05 significance level, 80% power, a group ratio of 1:1 and adjusting for 10.0% attrition rate and 10% non-compliance.

Figure 1: Baseline sample allocation by county and treatment status



3.1.3 Data collection

The midline data collection took place between February and March 2019. This period corresponds to exactly one year after the baseline stage of the study, which took place right before the tree kit and training intervention were delivered in 2019.

The midline stage of the study consisted of a quantitative survey administered to farmers that were interviewed at baseline. A tree trader survey was scheduled to take place immediately after the quantitative data collection, but Laterite decided to postpone this survey until teams can operate in the field again.

The midline survey was based on the baseline research instrument, with additional content related to the program implementation. We included additional modules designed to capture information on the tree value chain for the midline survey. The instrument covered household characteristics, socio-economic indicators, agricultural assets, and included a tree census of all the farmers' plots. Enumerators followed the same protocols set up at baseline for consistency.

3.1.4 Attrition

Measures put in place to minimize the risk of attrition included collecting multiple household contacts and GPS locations at baseline. This information aided in tracking participants during the midline survey. Attrition in this context is defined by a:

1. Participant who is completely unreachable via phone and cannot be traced using the GPS data collected at baseline.
2. Participant relocating from the study counties, hence not available for the survey.
3. Participant refusing to participate in the study, because they are no longer members of 1AF and wish not to be contacted again.

The attrition rate within the one-year follow-up was 7.3% in the control group and 5.9% in the treatment group. This falls within the attrition margin accounted for in the power calculations. While there was no association between treatment allocation and attrition (P-value = 0.225), attrition rate in Kericho County (7.8%) was significantly higher (P-value = 0.013) than Uasin Gishu (4.8%), at 5% significance level.

We tested whether there were any systematic differences between households that consented to the midline survey and those that could not be interviewed. We did this by studying joint significance; but the sample size is too small to conclude with confidence that there are no systematic differences between the two groups.

4 Findings

4.1 Demographics

The socio-demographic profile of respondents provides context on the study population. Of the 1,852 farmers interviewed at baseline, we completed interviews with 1,730 (93.4%). Of these, 92.7% from the control group and 94.1% from the treatment group were interviewed. The main reasons for the non-response/missing interviews were failure to consent, farmer opting out of 1AF program, relocation from the study counties, and inability to reach the farmer on phone or at their residence. The analysis focuses on the 1,730 farmers with complete data at baseline and midline.

While the data collection protocol specified that baseline farmers should be interviewed whenever possible, enumerators were allowed to interview a different household member with approval of their supervisor. We had 135 interviews, 72 in control group and 63 in the treatment group, where a different respondent was interviewed from the sampled household. This was mainly because the baseline respondent was unavailable at the time of the study. In scenarios like these, the enumerators conducted the interview with the member of the household most knowledgeable about 1AF activities.

Female farmers constituted more than half (59%) of the respondents and the mean farmer age was 45 years at midline. The participant age ranged from 18 to 89 years with a median of 43 years. While the difference on the average age between the control and treatment groups is not statistically significant, we observe that a higher proportion of control farmers are between 18 and 39 years of age compared to the treatment group (Table 1).

We describe the respondents' household income over the past 30 days preceding the survey as a proxy measure of their economic status. Household income was defined as the total earnings through wages, salaries, and self-employment earnings - that includes sale of trees, agricultural products, etc. in the past 30 days. We note that this is a self-reported measure and is subject to recall bias and to measurement error. Not all farmers were able to respond to this question.

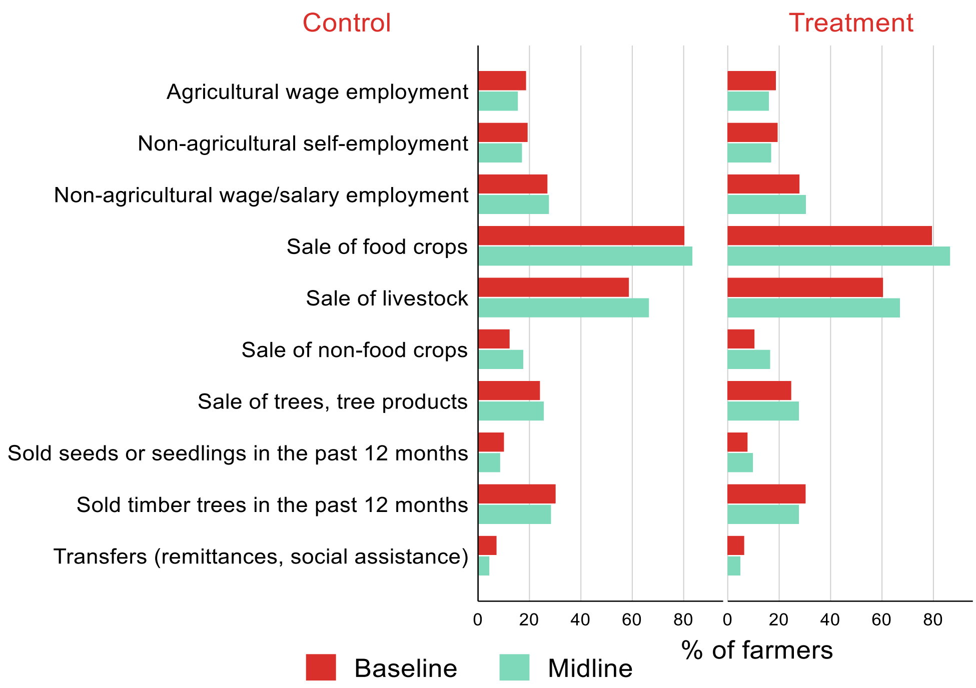
At midline, the average household income was 15,437 Ksh with a median income of 8,500 Ksh. The difference in income medians is statistically significant at the 1% level, with the control group having a median income of 9,000 Ksh. and the treatment group a median income of 8,000 Ksh. However, when we create income categories roughly representing the quintile distribution of the variable to account for the noise in the data, we find that there is no statistically significant difference in the proportion of farmers in each income category (Table 1).

Table 1: Midline participants' sociodemographic profile

	Control	Treatment	All	F/T-statistic (p-value)
Consent at midline				
Yes	92.7%	94.1%	93.5%	**1.39 (0.165)**
No	7.3%	5.9%	6.5%	
Total	**933**	**919**	**1,852**	
Gender				
Male	42.0%	40.0%	40.9%	**0.99 (0.321)**
Female	58.0%	60.0%	59.1%	
Total	**865**	**865**	**1,730**	
Age in years				
18-29	11.8%	10.1%	10.9%	**4.14 (0.001) *****
30-39	23.2%	28.6%	26.0%	
40-49	31.7%	26.7%	29.1%	
50-59	16.1%	19.4%	17.9%	
60-69	12.7%	10.1%	11.3%	
≥70	4.4%	5.2%	4.9%	
Mean age	45.2	45.1	45.1	**-0.27 (0.785)**
Total	**858**	**860**	**1,718**	
Median HH Income	9,000	8,000	8,500	**87.51 (0.000) *****
Income in Ksh.				**0.70 (0.59)**
0-2,999	16.0%	17.5%	16.8%	
3,000-4,999	12.3%	13.0%	12.7%	
5,000-9,999	22.6%	21.3%	21.8%	
10,000-19,999	20.4%	21.7%	21.1%	
≥20,000	28.7%	26.6%	27.6%	

The sale of crops and livestock (Figure 2) remain the most common income generating activities in both surveys and arms. There was notable increase in the proportion of farmers earning a living from the sales of food crops, livestock, non-food crops and trees, and tree products.

Figure 2: Income generating activities that farmers engaged in at baseline and midline by treatment status



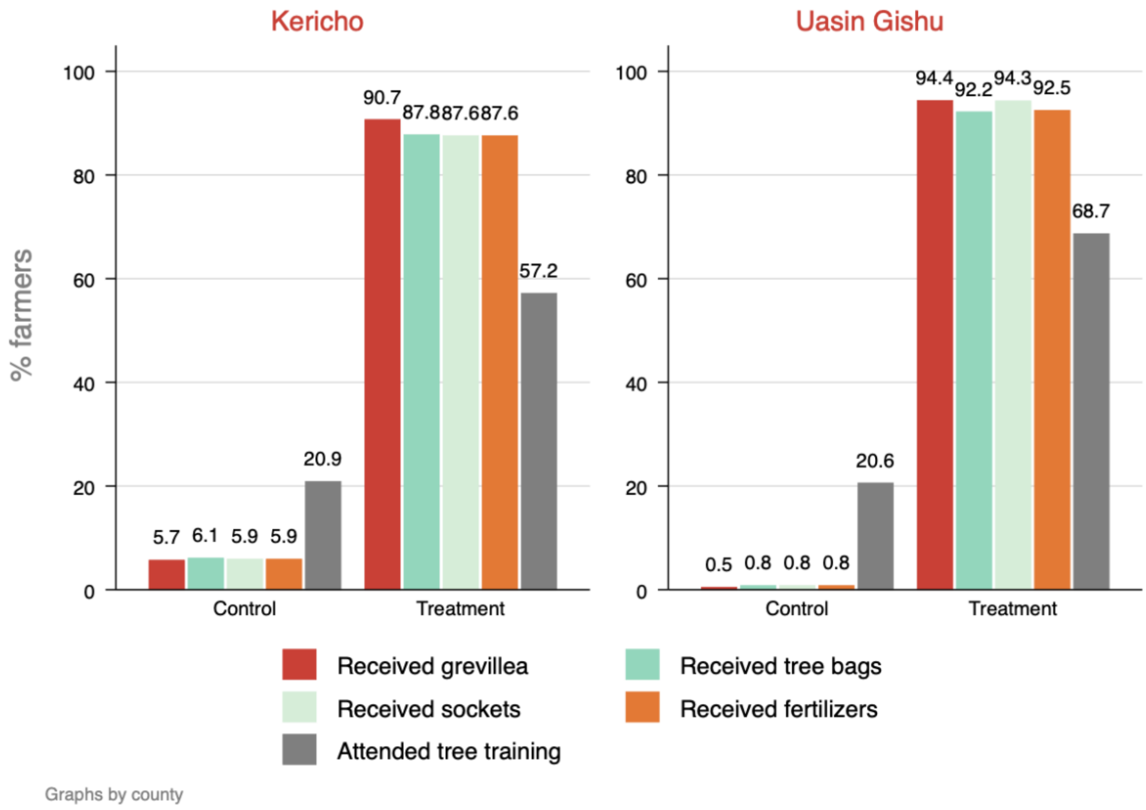
4.2 Treatment uptake

According to the protocol, a one-time tree intervention package was to be delivered in early/mid 2019, before the long rains, to farmers assigned to the treatment arm, after baseline data collection. The intervention was to be withheld from the control arm farmers until the end of the study. We explore the protocol compliance in the context of treatment uptake.

The study found (Table 2) that 3.4% control arm farmers received the grevillea tree kit and 7.8% of treatment farmers did not receive the tree kit. Self-reported attendance to the tree training shows that 20.8% of the control farmers attended the training on tree planting and 61.7% of the treatment farmers attended this training**.** According to 1AF, the delay in the 2019 long rains affected the training schedule and this caused the training to be rushed or not delivered at all in some areas. We discuss these findings further under the protocol compliance section, where we explain that the non-compliance rates are accounted for in the pre-analysis plan.

We find that full non-compliance rates, defined as receiving the trees and taking the specific tree training are low. While 92.2% of the treatment farmers received the tree kit, only 61.3% received the tree and attended the accompanying training. For the control farmers, 78.0% did not have any exposure to the program, meaning they did neither receive the tree kit nor the tree training.

Figure 3: Tree Program intervention uptake by treatment arm and county



We find that there are differences in the treatment uptake by county (Figure 4). Kericho reported more control group farmers (5.7%)[3] receiving the kit compared to Uasin Gishu (0.5%). The proportion of treatment farmers that failed to attend the tree training was also greater in Kericho, 42.8% compared to Uasin Gishu, 31.3%.

We note that this data needs to be contextualized. One Acre Fund was aware of the trainings being rushed because of the late long rains in 2019 and pointed out that some trainings might not have happened at all and that farmers might have planted the seeds without the training.

[3] 1AF had identified after baseline that trees had been distributed to some control farmers in Kericho.

Table 2: 1AF treatment uptake in the 2019 long rains season

	Kericho		Uasin Gishu		All counties	
	C	T	C	T	C	T
Received grevillea tree kit (seeds)						
Yes	5.7%	90.8%	0.5%	93.8%	3.4%	92.0%
No	94.2%	9.3%	99.5%	5.6%	96.5%	7.8%
Do not know	0.2%	0.0%	0.0%	0.6%	0.1%	0.2%
No. of observations	**491**	**526**	**374**	**339**	**865**	**865**
Received the tree bag with grevillea seeds						
Yes	6.1%	87.3%	0.1%	90.9%	3.8%	88.7%
No	93.5%	12.2%	99.9%	7.7%	96.0%	10.4%
Do not know	0.4%	0.5%	0.0%	1.4%	2.3%	0.9%
No. of observations	**491**	**526**	**374**	**339**	**865**	**865**
Received tree sockets with grevillea seeds						
Yes	5.9%	87.3%	0.8%	93.0%	3.7%	89.5%
No	93.7%	12.3%	99.2%	5.6%	96.1%	9.7%
Do not know	0.4%	0.4%	0.0%	1.5%	0.2%	0.8%
No. of observations	**491**	**526**	**374**	**339**	**865**	**865**
Received fertilizer with grevillea seeds						
Yes	5.9%	86.9%	0.8%	91.2%	3.7%	88.6%
No	93.7%	12.3%	99.2%	7.4%	96.1%	10.4%
Do not know	0.4%	0.8%	0.0%	1.5%	0.2%	1.0%
No. of observations	**491**	**526**	**374**	**339**	**865**	**865**
Attended agricultural training in the past 12 months?						
Yes	85.2%	87.4%	89.3%	90.0%	87.0%	88.4%
No	14.8%	12.6%	10.7%	10.0%	13.0%	11.6%
No. of observations	**491**	**526**	**374**	**339**	**865**	**865**
Attended tree planting training						
Yes	20.9%	57.2%	20.6%	68.7%	20.8%	61.7%
No	79.1%	42.8%	79.4%	31.3%	79.2%	38.3%
No. of observations	**491**	**526**	**374**	**339**	**865**	**865**

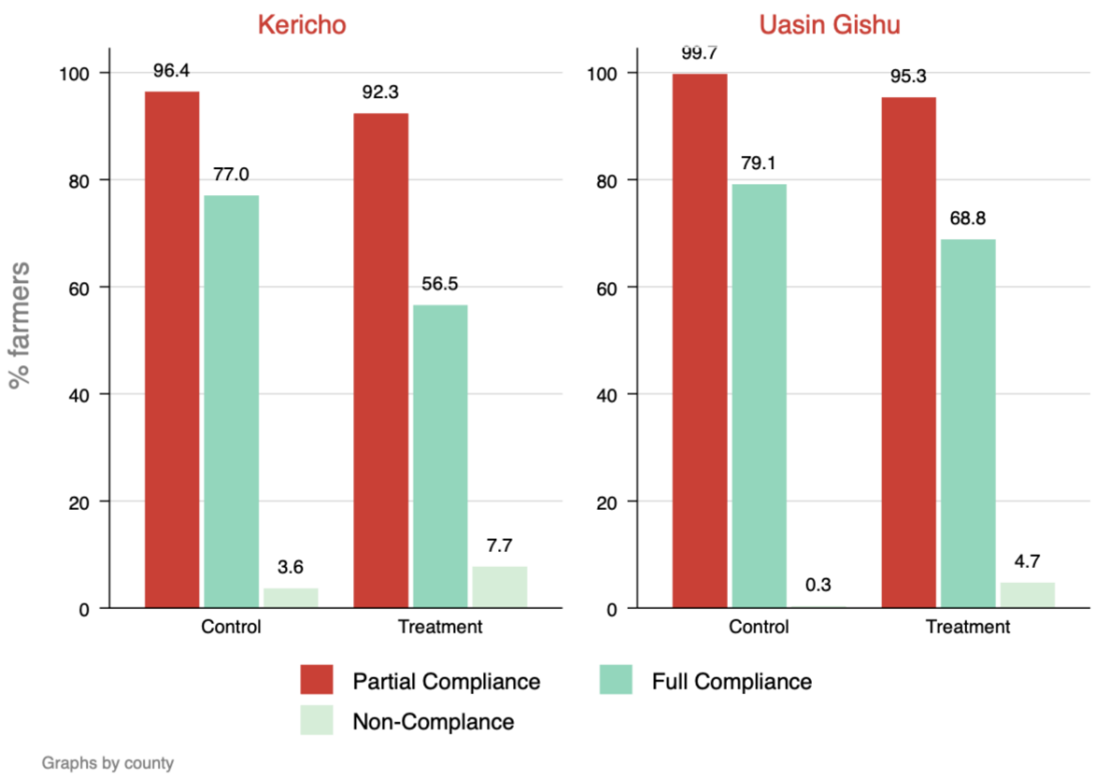
4.2.1 Protocol compliance

In order to understand what explains the observed level of non-compliance, we describe the characteristics of compliers and non-compliers. Three levels of compliance were defined as follows:

- **Fully compliant** – a treatment farmer who received the grevillea tree kit and attended the tree training or a control farmer who did not receive the kit and did not attend the tree training.
- **Partially compliant –** a treatment farmer either received the kit but failed to attend the training or did not receive the kit but attended the training; a control farmer who did not both receive the tree kit and the tree training.
- **Non-compliant** – a treatment farmer who did not receive the kit and failed to attend the training, or a control farmer who received the kit and attended the tree training.

We find that over 90% of farmers are partially compliant across treatment arms and counties (Figure 4). The overall partial compliance rate is 97.9% for the control group and 93.4% for the treatment group. Full compliance is much lower at 77.9% for the control group and 61.3% for the treatment group. This difference is explained by the low self-reported attendance rates in the treatment group and the control farmers that attended tree training.

Figure 4: Protocol compliance by treatment arm and county



We explore whether there are underlying characteristics associated to the two levels of compliance. For this analysis we focus on the treatment group and highlight that the high partial compliance rate leads to low statistical power for that analysis.

We run a logit regression with partial- and full- compliance as the independent variables and baseline demographic and socio-economic characteristics as the covariates to look for significant associations. We look at the farmer's county of residence, gender, age, education level, number of household members, 1AF contract value, number of income-generating activities in the past 12 months before baseline, and tree planting patterns at baseline.

We find that there is joint positive and significant association between the covariates and full compliance for the treatment group at the 1%. Living in Uasin Gishu, having planted grevillea trees in the 12 months before baseline, and participating in a higher number of income-generating activities in the 12 months before the baseline are positively correlated with being fully compliant (these associations are significant at the 5% level). Albeit weakly significant, already having any grevillea trees at baseline was also negatively associated with taking the up the full treatment (grevillea seeds and training). Additionally, demographic characteristics such as farmer gender and age are not significantly correlated with full protocol compliance.

For partial compliance, we are limited by statistical power because of the high proportion of partial compliers.

Table 3: Determinants of protocol compliance

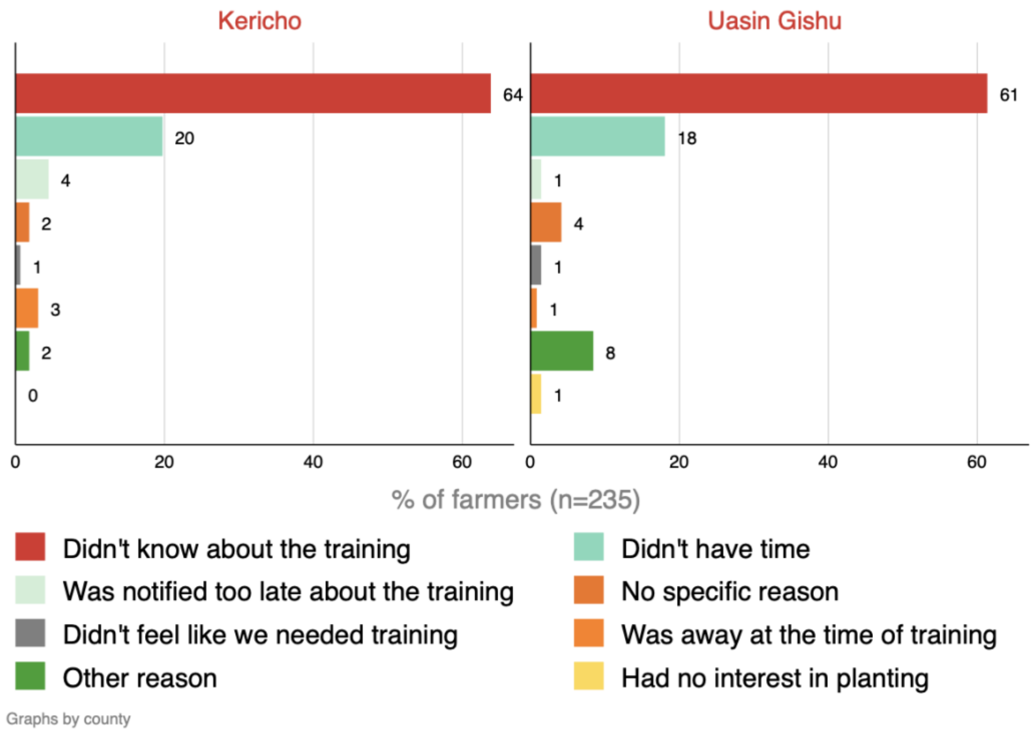
Variable	Full Compliance	
	Coefficient [Std. Error]	P>|z|
County (Uasin Gishu)	.4853*** [.1787]	0.007
Share of 1AF farmers that planted grevillea trees in the past 12 months (2019)	.6172** [.2837]	0.030
Share of 1AF farmers with any grevillea trees (2019)	-.3471* [.1827]	0.057
Number of grevillea trees on land (2019)	-.0012 [.0019]	0.544
Share of 1AF farmers that planted any trees in the past 12 months (2019)	.1824 [.1469]	0.214
Household size (2019)	.0048 [.0288]	0.869
Self-reported land area in acres (2019)	-.0088 [.0098]	0.369
Participant gender (2019)	.1921 [.1499]	0.200
Participant age (2019)	.0291 [.0337]	0.389
Participant age^2 (2019)	-.0003 [.0003]	0.382
Share of participants that completed primary education (2019)	.2571 [.1656]	0.510
Share of participants that planted maize as a major crop (2019)	.0219 [.1949]	0.910
Number of income generating activities in the past 12 months (2019)	.1501** [.0763]	0.049

4.2.2 Attendance of tree training

While we are aware that it is possible that not all tree trainings took place, we explore the farmers' self-reported tree training attendance. We further look at curriculum specific questions asked to assess whether participating farmers increased their tree planting knowledge through the training. For this analysis, we focus on farmers in the treatment group.

We learn that the most common reason for not attending the tree training is not knowing about it (Figure 5). Overall, 63.0% of the farmers in the treatment group that did not attend the tree training did not know about it. The rate is slightly higher for Kericho than for Uasin Gishu, but this difference is not statistically significant. Additionally, a fifth (19.2%) of the farmers state not having time to attend the tree training as a reason not to attend.

Figure 5: Reasons for not attending the tree training



We further explore whether the self-reported attendance to the tree training is correlated to increased knowledge on the topics taught in the training. For this, we compare the knowledge of treatment farmers that attended and those that did not attend the tree training on questions extracted from the Grevillea Tree Kit Training Booklet (One Acre Fund, 2018).

We find that treatment farmers that self-reported attendance to tree training tend to have better knowledge of tree planting best practices compared to farmers who did not attended the training. A higher proportion of farmers that report attending the training know the correct soil for planting, the time they should wait for a grevillea seed to germinate, how often they should water grevillea seeds, best practices to maximize seed survival. The farmers that self-report attending the tree training have also an overall higher score on the number of knowledge questions they respond correctly, compared to farmers that state not to have attended. Farmers that report not attending the training are more knowledgeable about pruning 25% or less of the tree branches than farmers that report attending the training.

Table 4: Grevillea planting knowledge by self-reported training attendance

Knowledge	Not Trained (N=332)	Trained (N=533)	P-Value
Correct Soil	38%	57%	0.000***
Correct Germination	4%	7%	0.050**
Correct Watering	87%	92%	0.028**
Correct Survival	96%	99%	0.019**
Correct Spacing	83%	85%	0.497
Correct Pruning (Proportion)	61%	52%	0.003***
Correct Socketing	12%	10%	0.492
Correct Transfer	49%	49%	0.938
Correct Hardening	2%	1%	0.216
Correct Planting	14%	14%	0.906
Correct Pruning (Timing)	35%	35%	0.946
Knowledge Index (0-11)	4.7	4.9	0.021**

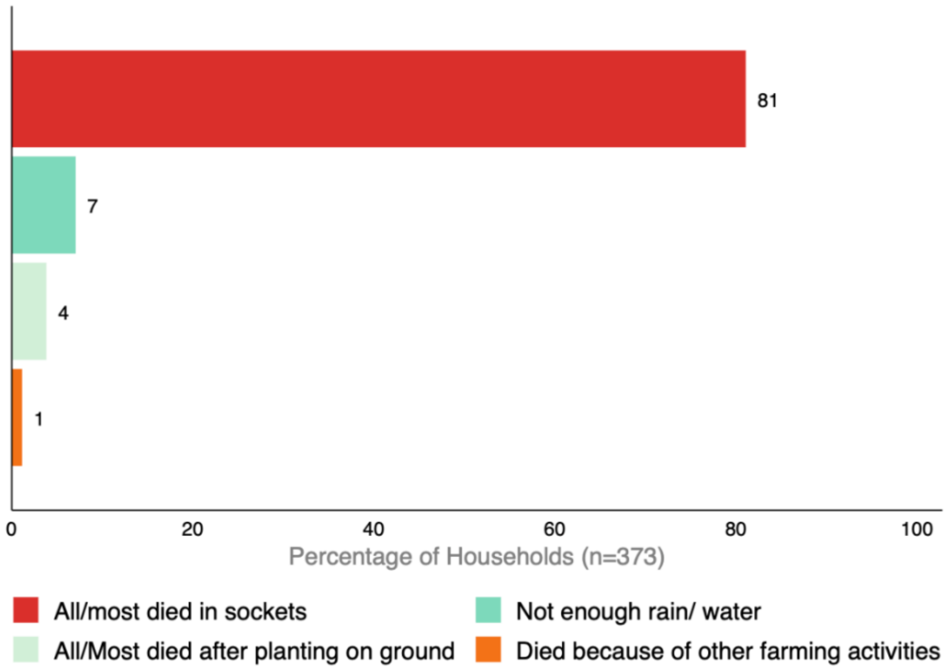
4.3 Grevillea tree planting and survival

The evaluation seeks to investigate whether treatment farmers have additional trees on average and how does this translate to in terms of potential financial value by the end of the study.

Overall, 68.7% of the treatment farmers planted grevillea seeds from the tree kit in tree bags. The planting rate for the grevillea seeds was 65.8% in Kericho and 73.2% in Uasin Gishu. For the control group, 3.0% of the farmers planted trees from the grevillea tree kit. Most of the treatment farmers (63.8%) that received the grevillea seeds and did not plant them, stated that they would plant the seeds at a later date.

However, only a quarter (24.9%) of treatment farmers planted tree seedlings on the ground. Overall, 21.2% of the treatment farmers claimed to have grevillea trees planted on their plots from the tree kit. The main reason for trees not surviving is that the seedlings died in the sockets (Figure 6). This is followed by lack of rain, and the trees dying when planted on the ground.

Figure 6: Reasons for not planting trees on the ground



On average, treatment farmers planted 59 trees on the ground with a median of 20 trees. The average number of trees that survived was 32 with a median of 15 trees. The large difference between the mean and median is due to a few farmers that have 200 or more grevillea seeds surviving.

The mean survival rate in the treatment group for trees planted on the ground was 66.7% at midline. This means that two-thirds of all seedlings from the tree kit that were planted on the ground had survived at midline, according to the farmers.

Table 5: Uptake of tree planting activities among treatment farmers by county of residence

Variable	Kericho	Uasin Gishu	All
Planted any grevillea seeds from the tree kit			
Yes	65.8%	73.2%	68.7%
No	34.2%	26.8%	31.3%
No. of observations	**526**	**339**	**865**
Planted any grevillea seedlings from the tree kit on the ground			
Yes	22.9%	28.1%	24.9%
No	77.1%	71.9%	75.1%
No. of observations	**526**	**339**	**865**
Number of grevillea seedlings planted on the ground	67	48	59
No. of observations	**120**	**95**	**215**
Had grevillea trees from the tree kit surviving at midline			
Yes	18.9%	24.8%	21.2%
No	81.1%	75.2%	78.8%
No. of observations	**526**	**339**	**865**
Number of grevillea trees from the tree kit surviving	31	34	32
No. of observations	**99**	**84**	**183**
Survival rate of grevillea trees from the tree kit (#surviving / #planted)	65.2%	68.8%	66.7%
No. of observations	**120**	**95**	**215**

In addition to the main impact evaluation questions, we sought to investigate whether there is a correlation between self-reported attendance to the tree training and grevillea tree survival rates. The survival rate is defined as the number of grevillea trees from the tree kit surviving divided by the number of grevillea seeds planted on the ground. We note that his analysis is limited by statistical power as only a quarter of treatment farmers planted seedlings from the kit on the ground.

We do not find a statistically significant correlation between attending the tree training and the survival rate for grevillea trees from the tree kit. While the coefficient from this regression is negative, it is not statistically significant. We also do not find any other demographic and socio-economic variable from baseline to be correlated to tree survival rates at midline.

We do find a positive and statistically significant (at the 1%) level association between knowledge on tree planting best practices and survival rates for the treatment group. Farmers that responded correctly to more knowledge questions are more likely to have surviving grevillea trees on their plots.

Finally, we explore whether treatment farmers that i) planted any grevillea seeds from the tree kit and ii) planted and grevillea seedlings from the tree kit on the ground have different baseline characteristics than those who did not. We look at the farmers' gender, age education level, household size, number of income activities, and ownership of grevillea, timber and overall trees at baseline for this comparison.

We find that farmers that planted seedlings on the ground were significantly more likely to own more (timber-) trees at baseline that those who did not. A higher proportion of the farmers that planted seedlings completed primary education and these farmers were engaged in a marginally higher number of income activities at baseline compared to the farmers that did not plant the seedlings. We also studied the joint significance of the baseline variables and find that the difference between both groups is statistically significant at the 1% level. However, the sample size is too small to conclude with confidence that there are systematic differences between the two groups.

Table 6: Baseline characteristics of treatment farmers that planted grevillea seedlings

Variable	No Seedlings Planted (N=690)	Seedlings Planted (N=214)	P-Value
Respondent Gender	58%	62%	0.26
Respondent Age	44	44	0.81
Completed Primary	68%	74%	0.09*
Household Members	5	5	0.84
Number of Income Activities	2.4	2.6	0.02**
Grevillea Ownership (Y/N)	48%	45%	0.52
50 Timber Trees or More (Y/N)	51%	70%	0.00***
100 Trees or More (Y/N)	45%	60%	0.00***

The results for the farmers that planted seeds from the tree kit point in the same direction, with a higher proportion of those that planted the seeds owning a higher number of (timber-) trees at baseline. For this comparison, a higher proportion of farmers is female compared to those that did not plant the seeds and the farmers that planted the tree kit seeds were also marginally engaged in more income generating activities at baseline. The joint significance of the baseline variables is statistically significant at the 5% level. For simplicity, we present the results of the comparison between farmers that planted seeds from the tree kit and those that did not on Appendix 1.

4.4 Treatment impact on grevillea trees

Next, we look at the impact of the Tree Program on the planting of grevillea trees. For this, we look at whether the farmer planted any grevillea trees in the past twelve months and at the number of grevillea trees planted. We use multiple analysis methods to check the robustness of our estimates. First, we run a simple difference in difference regression on both planting rates and number of grevillea trees planted in the past twelve months. Next, we add socio-economic controls using baseline values to check whether the coefficients remain significant with controls. Further, for the number of grevillea trees planted, we run the same regressions, using the natural logarithm plus one to eliminate issues with the non-normal distribution of the number of trees planted that we observe. Finally, we run a regression using the difference between the number of grevillea trees at midline and baseline, controlling for treatment status and the number of grevillea trees planted at baseline[4].

Table 7: Summary of grevillea planting at baseline and midline

	Baseline		Midline	
	Control	**Treatment**	**Control**	**Treatment**
Planted grevillea trees in the past 12 months	8.8%	8.1%	10.5%	25.3%
No. of observations	**933**	**919**	**864**	**864**
Number of grevillea trees planted in the past 12 months	2.4	1.3	1.6	8.2
No. of observations	**933**	**919**	**864**	**864**

Overall, we find a consistent positive effect of the Tree Program on grevillea planting (Table 8). The results show that the program increased the share of farmers that planted grevillea trees in the past 12 months by 15.6 percentage points, holding everything else equal, using the difference in difference model. This coefficient is statistically significant at the 1% level both in the simple model and in the model with controls. Control variables include county of residence, household size, and the gender, age and education of the participant. In the difference in difference model, the coefficients for the county of residence (residing in Uasin Gishu), and being female are negative and significant at the 1%. The coefficient for the age of the farmer is positive and significant at the 10% level, but this effect diminishes as farmers get older. Finally, the coefficient for whether the farmer completed primary education is positive and significant at the 10% level.

The results for the lag model are similar to the difference in difference model and are also statistically significant at the 1% level. This applies to the simple specification and for the model with controls. For this model, the coefficients of control variables take the same sign as in the difference in differences model, but only county of residence and gender of the farmer are significant, at the 10% level. The higher R-Squared coefficient for the lagged regressions shows that the lag is strongly predictive of the final outcome. I.e. Whether a farmer had planted grevillea trees before the start of the program (at baseline) is strongly predictive on whether they planted grevillea trees at the midline stage. The higher R-Squared for this model would indicate that the estimates of the lag with a difference model would be more accurate than other models. However, in this case we find that coefficients are very close to each other, independent of the model chosen.

[4] This type of regression is called a difference with a lag. It also produces an estimate of the treatment effect. It is an alternative to difference-in-difference estimates, when the core assumption of parallel trends does not hold. This regression used baseline values to estimate whether being part of the treatment group has an effect on the difference of the outcome variable between midline and baseline.

Table 8: Impact of the program on grevillea planting rates

Dependent variable	Model	Coefficient [Std. Error]	P-Value	R^2
Planted grevillea trees in the past twelve months (Yes/No)	Simple difference in differences	.1557*** [.0172]	0.000	0.0447
	Difference in differences with controls	.1566*** [.0173]	0.000	0.0617
	Simple lag	.1507*** [.0151]	0.000	0.2435
	Lag with controls	.1517*** [.0153]	0.000	0.2488
No. of observations			**1,728**	

While the objective of this evaluation is to measure the intention to treat (ITT) effect of the program, we also estimate the average treatment effect on the treated, referred to as the local average treatment effect (LATE). For this, we use instrumental variable (IV) estimations for both farmers that are partially and fully treated. For this estimation we use whether a farmer is partially or fully compliant with the treatment as the IV. The underlying assumption for this model is that compliance is highly correlated with the treatment status of the farmers and that compliance only affects whether farmers planted grevillea trees in the past 12 months through the treatment status of the farmer.

We find that the treatment effect is higher for fully compliant farmers than for partially compliant farmers, which in turn is higher than the intention to treat effect. This can be seen in Table 9, which shows an average treatment effect of 22.1 percentage points for the partially treated and 27.0 percentage points for the fully treated.

Table 9: Local average treatment effect on grevillea planting rates

Dependent variable	Model	Coefficient [Std. Error]	P-Value	R^2
Planted grevillea trees in the past twelve months (Yes/No)	IV Regression for Partial Compliers	.2197*** [.0243]	0.000	0.0164
	IV Regression for Full Compliers	.2688*** [.0300]	0.000	0.0034
No. of observations			**1,728**	

For the number of trees planted in the past 12 months, we also observe a positive and statistically significant effect for the three models tested. The difference in difference model shows that the treatment is associated with an average increase of 7 grevillea trees planted, holding everything else equal. This result is significant at the 1% level. We also find similar patterns with the control variables, whereby the coefficients for living in Uasin Gishu and being female are negative, while age and education are positive, with the former diminishing as farmers get older. In the difference in differences model, only county and gender are statistically significant at the 10% level. We report the results from the difference in differences regression with controls for simplicity and because models are consistent across the board. We report other model specifications Appendix 2.

Table 10: Impact of the program on grevillea tree numbers

Dependent variable	Model	Coefficient [Std. Error]	P-Value	R^2
Number of grevillea trees planted in the last 12 months	Difference in differences with controls	7.5448*** [1.3793]	0.000	0.0169
No. of observations			**1,728**	

We find a higher treatment effect for partially- and fully compliant farmers on the number of grevillea trees planted in the last 12 months, compared to the analysis for the treatment group only. The local average treatment effect regression shows that partially compliant farmers planted an average of 11 grevillea trees, while fully compliant farmers planted an average of 13 trees. This compares to the 7 grevillea trees from the difference in differences estimate for the intention to treat effect. Both LATE coefficients are significant at the 1% level.

We further test the local average treatment effect on the number of grevillea trees for i) treatment farmers that planted any seeds from the tree kit, and ii) treatment farmers that planted grevillea seedlings that germinated from the planted tree kit seeds. We find that farmers that planted at least one seed from the tree kit have a higher local average treatment effect compared to partially compliant farmers (those that either received the tree kit or the tree planting training), but a lower LATE than fully compliant farmers (those that received both the tree kit and the tree planting training). We find the highest LATE for treatment farmers that planted at least one germinated seedling from the tree kit on the ground. The local average treatment effect for this group shows an increase of 31 grevillea trees and is significant at the 1% level.

Table 11: Local average treatment effect on grevillea tree numbers.

Dependent variable	Model	Coefficient [Std. Error]	P-Value	R^2
Number of grevillea trees planted in the last 12 months	IV Regression for Partial Compliers	10.9283*** [1.9828]	0.000	0.0074
	IV Regression for Full Compliers	13.3727*** [2.4235]	0.000	0.0074
	IV Regression for Planted Seeds	11.3456*** [2.0430]	0.000	0.0206
	IV Regression for Seedlings Planted on the Ground	31.2984*** [5.5330]	0.000	0.0555
No. of observations			**1,728**	

4.5 Treatment impact on tree assets

We explore whether the treatment has any effect on the overall number of trees planted in the last 12 months by the farmer and in the allocation of trees planted in this period of time. For this, we look at the absolute number of (timber) trees planted by farmers in the 12 months preceding the data collection, as well as the grevillea trees planted as a percentage of (timber-) trees planted in the last 12 months. We also look at non-grevillea trees as a percentage of timber trees planted in the last 12 months.

We observe that more farmers planted trees in the 12 months preceding the midline survey compared to the 12 months before baseline in both the control and treatment groups, as shown on Table 12. This applies for both timber trees and fruit trees in particular. However, while the average number of timber trees planted in the past 12 months remains relatively constant between baseline and midline for both groups, there is an uptake in the average of number of fruit trees planted at midline for both the treatment and control groups compared to baseline.

Table 12: Summary of tree planting at baseline and midline

	Baseline		Midline	
	C	T	C	T
Planted timber trees in the past 12 months	43.1%	42.8%	50.6%	51.8%
No. of observations	**865**	**865**	**865**	**865**
Planted fruit trees in the past 12 months	31.8%	31.4%	44.4%	46.4%
No. of observations	**865**	**864**	**865**	**865**
Planted other trees in the past 12 months	12.5%	13.4%	14.2%	17.2%
No. of observations	**865**	**865**	**865**	**865**
Planted all trees in the past 12 months	60.3%	60.8%	69.8%	70.5%
No. of observations	**865**	**864**	**865**	**865**
Number of timber trees planted in the past 12 months	31.6	33.3	30.0	35.5
No. of observations	**865**	**865**	**865**	**865**
Number of fruit trees planted in the past 12 months	6.8	8.8	9.8	12.7
No. of observations	**865**	**864**	**865**	**865**
Number of other trees planted in the past 12 months	3.0	3.9	3.2	6.5
No. of observations	**865**	**865**	**865**	**865**
Number of all trees planted in the past 12 months	45.0	48.3	45.4	62.3
No. of observations	**865**	**864**	**865**	**865**
Grevillea trees as % of timber trees planted In the past 12 months	5.2%	4.7%	5.6%	14.7%
No. of observations	**865**	**864**	**865**	**864**
Grevillea trees as % of all trees planted In the past 12 months	3.6%	3.4%	3.6%	11.0%
No. of observations	**865**	**865**	**865**	**865**
Non-grevillea trees as % of timber trees in the past 12 months	94.8%	95.3%	94.4%	85.3%
No. of observations	**865**	**864**	**865**	**864**
Non-grevillea timber trees as % of all trees in the past 12 months	30.5%	30.8%	35.0%	26.9%
No. of observations	865	865	865	865

We find an overall positive, but not statistically significant, treatment effect on whether the farmers planted any timber, fruit, or any trees in the past 12 months and on the number of trees from these categories planted. The results point towards the possibility of these effects accumulating when we look at the total number of trees planted in the past 12 months, where we see a large (~14 trees) effect that is statistically significant at the 10% level. However, we do not have enough statistical power to confirm this hypothesis and the positive coefficients on all effects could be due to random noise.

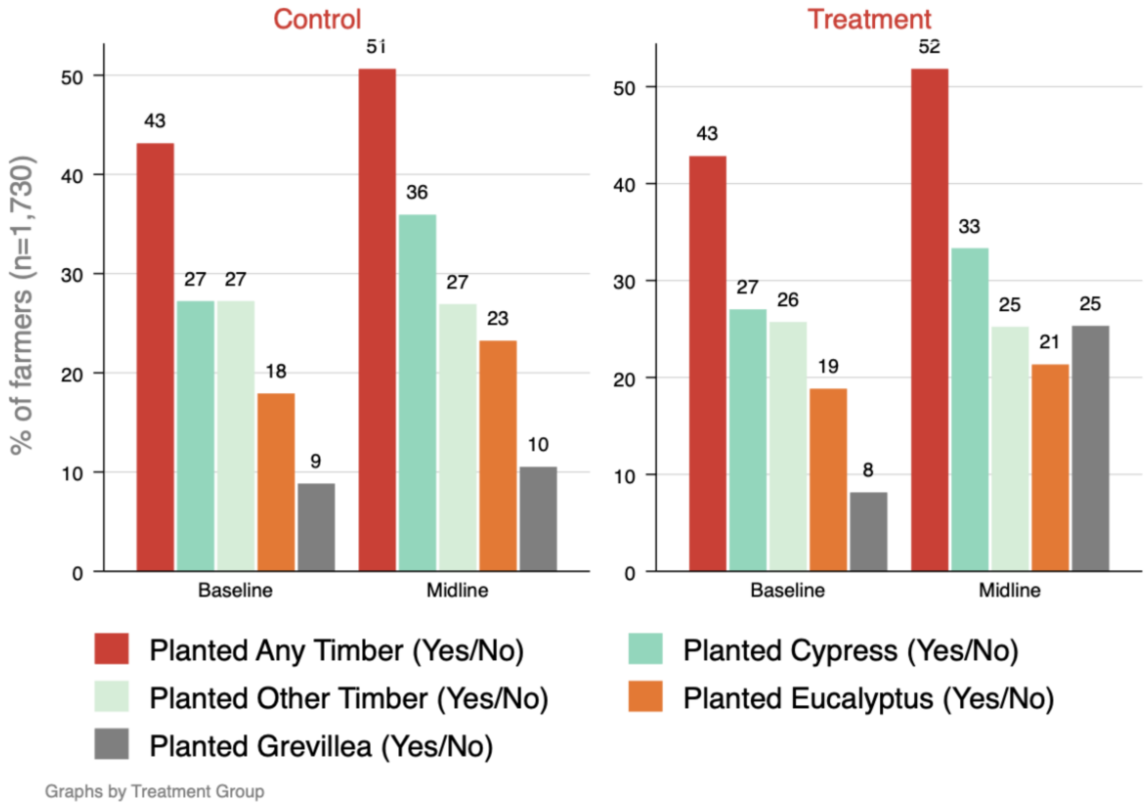
We also find that the Tree Program has a positive and significant effect on the number of young grevillea trees as percentage of all (timber) trees planted (Table 13). The effect of farmers replacing newly planted timber trees with grevillea is reiterated when we look at the proportion of non-grevillea timber trees as percentage of all (timber-_ trees planted in the last 12 months. For simplicity, we only report the results of the difference in difference model with controls. These results hold when applying lag regressions. We present the results on the average treatment effect on the treated in Appendix 3. The LATE results are aligned with the main results, whereby coefficients are higher for partial- and full compliers, respectively, compared to the average treatment effect. The significance levels of the LATE coefficients remain the same as in the main regressions.

Table 13: Impact of the program on tree planting

Dependent variable	Coefficient [Std. Error]	P-Value	R_2	No. of observations
Planted timber trees in the past 12 months	.0148 [.0263]	0.574	0.0219	1,730
Planted fruit trees in the past 12 months	0.0238 [.0238]	0.337	0.0346	1,730
Planted other trees in the past 12 months	.0209 [.0194]	0.173	0.0054	1,730
Planted any trees in the past 12 months	.0034 [.0246]	0.890	0.0219	1,730
Number of timber trees planted in the past 12 months	3.818 [6.319]	0.546	0.0061	1,730
Number of fruit trees planted in the past 12 months	1.0019 [2.3973]	0.676	0.0129	1,730
Number of other trees planted in the past 12 months	2.5254 [1.716]	0.143	0.0082	1,730
Number of trees planted in the past 12 months	13.7732* [7.7294]	0.076	0.0193	1,730
Grevillea trees as % of timber trees planted in the past 12 months	.0965*** .0125]	0.000	0.0510	1,730
Grevillea trees as % of all trees planted in the past 12 months	.0768*** [.0100]	0.000	0.0451	1,730
Non-grevillea timber trees as % of timber trees planted in the past 12 months	-.0964*** [.0125]	0.000	0.0510	1,730
Non-grevillea timber trees as % Of all tree planted	-.0850*** [.0215]	0.000	0.0198	1,730

We explain the lack of a measurable treatment effect on whether farmers planted timber trees in the past 12 months by the overall increase of planting of timber trees across both groups and by the fact the treatment group is substituting other timber trees with grevillea (Figure 7). In both the treatment and control groups a higher proportion of farmers had planted timber trees in the past twelve months at midline. While we observe a clear increase on the proportion of farmers planting grevillea in the treatment group, this increase is matched by farmers in the control group, where more farmers planted cypress and eucalyptus trees. These results hold when we look at the average treatment effect on the treated for whether the farmers planted timber trees in the past 12 months, the results for the farmers that are partially- or fully compliant with the treatment protocol are not statistically significant for this outcome.

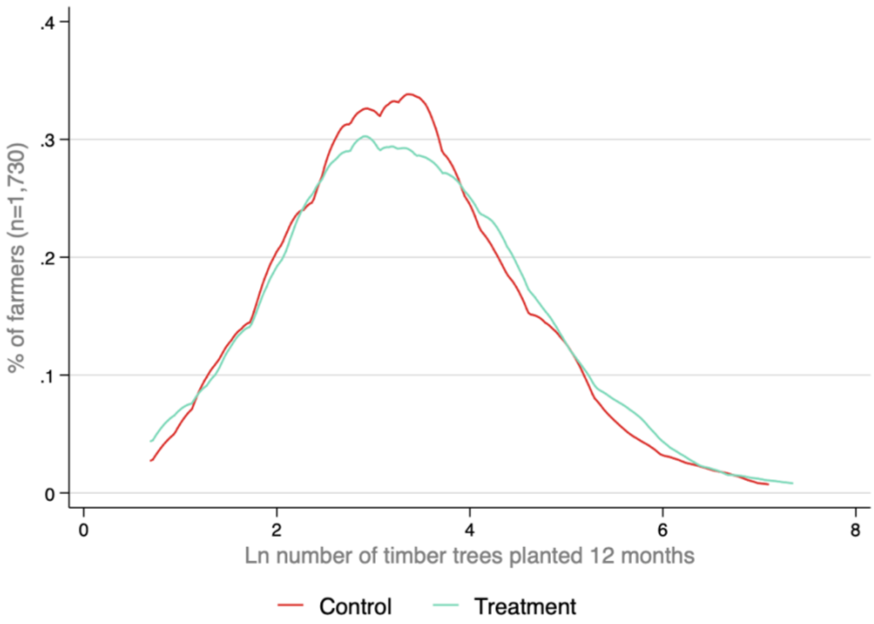
Figure 7: Timber tree planting rates in the past 12 months

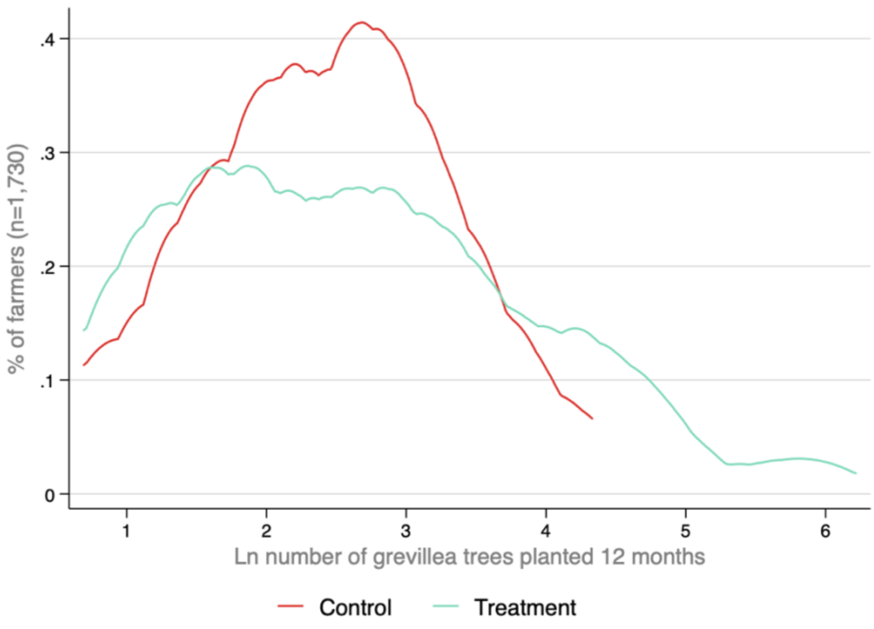


This effect cannot only be found when comparing treatment and control but also when comparing partial- and full compliers to non-compliers within the treatment group. Non-compliers experience a more rapid increase in the planting of cypress and eucalyptus trees, while partial and full compliers experience an increase in the planting of grevillea trees. Full compliers planted on average 21 non-grevillea timber trees in the 12 months before midline while those not fully complying planted 35 non-timber trees in the same period.

These results can be reiterated graphically on Figure 8. This figure shows that the distribution of timber trees between treatment and control groups is similar at midline, the treatment group clearly replaces newly planted timber trees with grevillea.

Figure 8: Ln timber and grevillea trees planted in the past 12 months





4.6 Heterogenous effects

Lastly, we explore how the treatment effect is different for the two different counties.

When we look at the effects of the treatment on the planting of grevillea trees, we find that these are greater in Uasin Gishu than in Kericho (Table 14). For all the variables where we found a positive and statistically significant impact of the program in the overall population, we find higher coefficients and a higher R-squared statistic for the regressions in Uasin Gishu.

Table 14: Impact of the program by county

Dependent variable	County	Coefficient [Std. Error]	P-Value	R_2	No. of observations
Planted grevillea trees in the past 12 months (Yes/No)	Kericho	.1223*** [.0247]	0.000	0.0379	1,008
	Uasin Gishu	.2084*** [.0235]	0.000	0.1085	711
Number of grevillea trees planted in the past 12 months	Kericho	7.5256*** [2.1040]	0.000	0.0142	1,008
	Uasin Gishu	8.337*** [1.8858]	0.000	0.0343	711
Planted timber trees in the past 12 months (Yes/No)	Kericho	-.0199 [.0338]	0.555	0.0214	1,009
	Uasin Gishu	.0677 [.0419]	0.111	0.0257	711
Number of timber trees planted in the past 12 months	Kericho	3.0894 [8.8627]	0.728	0.0044	1,009
	Uasin Gishu	5.6150 [8.721]	0.522	0.0130	711
Planted fruit trees in the past 12 months (Yes/ No)	Kericho	.02965 [.0326]	0.364	0.0295	1,009
	Uasin Gishu	.0181 [.0378]	0.633	0.0353	711
Number of fruit trees planted in the past 12 months	Kericho	2.3266 [1.8464]	0.210	0.0116	1,009
	Uasin Gishu	-.2915 [5.1377]	0.955	0.0108	711
Planted any trees in the past 12 months (Yes/No)	Kericho	-.0164 [.0314]	0.601	0.0241	1,009
	Uasin Gishu	.0323 [0396]	0.417	0.0206	711
Number of trees planted in the past 12 months	Kericho	5.0067 [9.9454]	0.615	0.0070	1,009
	Uasin Gishu	28.5750** [12.3717]	0.024	0.0288	711
Grevillea trees as % of timber trees planted in the past 12 months	Kericho	.0891*** [.0185]	0.000	0.0305	1,009
	Uasin Gishu	. 10924*** [.01536]	0.000	.01536	711
Grevillea trees as % of all trees planted in the past 12 months	Kericho	.0638*** [.0147]	0.000	0.0245	1,009
	Uasin Gishu	.0977*** [.0128]	0.000	0.0903	711

4.7 Land use for crop farming

To better understand the changes in land use crop farming, with the adoption of tree planting, we look at the difference in the number of plots with a major crop[5] on land between baseline and midline. We notice that about half of the farmers, recorded no change is the number of plots with a major crop. A similar pattern was observed in both treatment groups (Table 15). We observe that about 70% of the farmers in both treatment groups did not change the allocation of plots that they use for tree planting.

Table 15: Change in land use for crop farming and tree planting

	Control	Treatment
Overall difference in number of plots with major crops between baseline and midline		
Decrease (1-4 plots)	20.5%	20.4%
No change	53.2%	54.2%
Increase (1-8 plots)	26.3%	25.4%
No. of observations	**865**	**865**
Overall difference in number of plots with trees between baseline and midline		
Decrease (1-4 plots)	12.7%	13.7%
No change	70.7%	71.3%
Increase (1-8 plots)	16.6%	15.0%
No. of observations	**865**	**865**

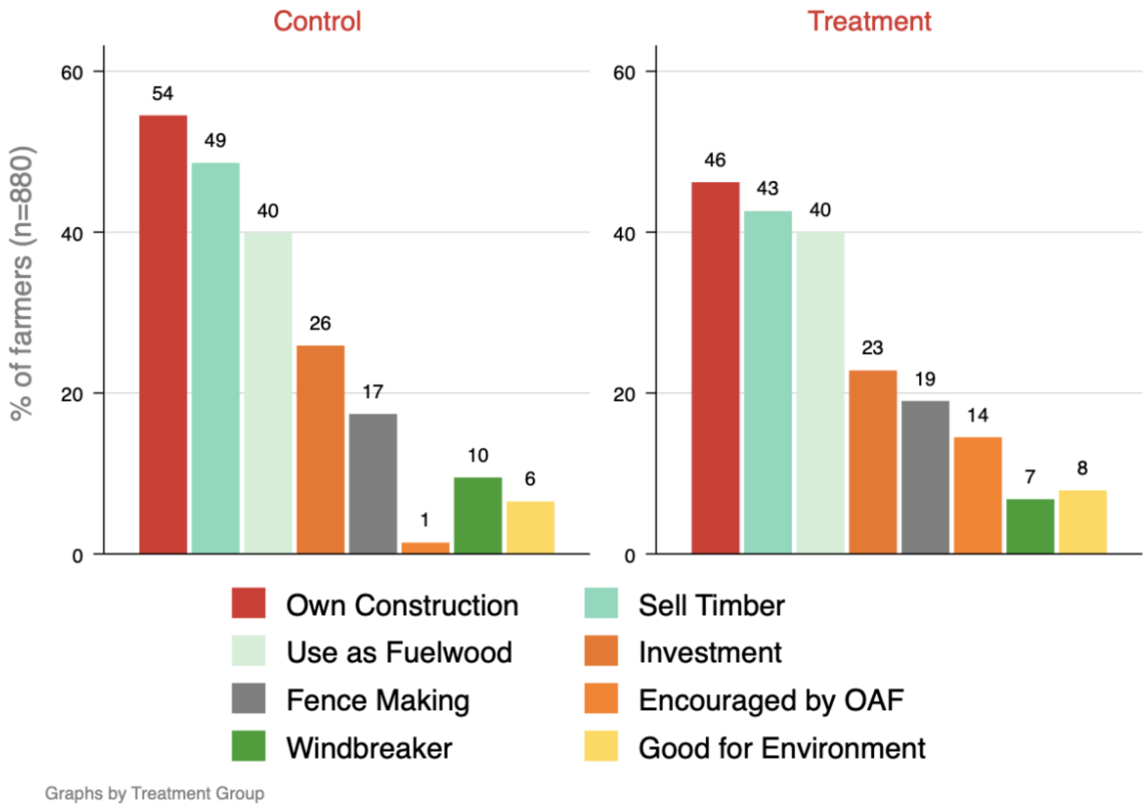
4.8 Tree planting perceptions

In this section, we explore the farmers' attitudes and perceptions towards tree planting. We focus on the reasons for planting timber trees and on the treatment group, while also observing the control group.

At midline, the most common reason for planting trees across both groups was own construction (Figure 9). This was followed by farmers wanting to sell the timber from the trees and to use the trees as fuelwood. Other reasons for planting timber trees included investments, making fences, use as windbreakers, and because trees are good for the environment.

[5] A crop type covering more than half of the plot.

Figure 9: Reasons for planting timber trees



While we do not ask directly why farmers planted grevillea trees, we look at the reasons for planting timber for treatment farmers that planted grevillea at midline. We find that the main three reasons for planting timber for this subsample are i) fuelwood for cooking (47.5%), ii) own construction (42.9%), and iii) encouragement by 1AF (25.4%).

4.9 Knowledge of tree-planting best practices at midline

This section highlights the changes in the farmers' knowledge of tree planting best practices. The questions are based on 1AF tree training modules.

We find a treatment effect for the knowledge on using the correct soil for planting tree seeds (Table 16). There was a 13.2 percentage point increase in the share of farmers that knew that they should gather soil from close to a healthy tree, use topsoil only, use a mix of topsoil and sand, or specifically not using clay soil at midline in the treatment group compared to the control group. This coefficient is statistically significant at the 1% level. In aggregate, the treatment also has a positive and statistically significant effect on the number of knowledge questions that the farmers answer correctly. There was no significant change in the knowledge of other best practices asked at baseline. These questions where about the time farmers should wait for a seed to germinate (two months or more), how often they should water the seeds if it does not rain (at least once a day), what they can do to maximize germination rates, and correct pruning (25% of the branches or less).

Table 16: Changes in knowledge of tree planting best practices

	Baseline		Midline		Difference in Difference
	Control	Treatment	Control	Treatment	
No. of observations	**865**	**865**	**865**	**865**	
Correct Soil	23.8%	24.5%	35.7%	49.8%	13.3***
Correct Germination	6.1%	4.8%	5.5%	5.6%	1.9
Correct Watering	85.5%	85.9%	89.4%	89.9%	0.8
Correct Survival	95.4%	94.4%	97.3%	97.6%	1.4
Correct Pruning	50.7%	48.0%	56.9%	55.4%	0.2
Knowledge Index (0-5)	2.6	2.6	2.8	3.0	0.2***

5 Key midline findings

The main findings from the midline correspond to the treatment effect on grevillea and tree planting. Other main findings correspond to the attrition and compliance rates, as well as the change in knowledge on tree planting best practices.

First, we find that the treatment had a statistically significant effect on whether the farmers planted grevillea trees in the past 12 months and on the number of grevillea trees planted by farmers in this time period. Using a difference in differences approach, we observe a 15.6 percentage point increase on whether the farmer planted grevillea trees in the past 12 months. For the number of trees, we find an increase of 7 grevillea trees. The results remain significant after various robustness checks.

Second, we find that while there is no overall treatment effect on whether farmers planted timber, fruit or other trees in the past 12 months and in the numbers of trees in these categories planted, there is a positive and statistically significant treatment effect on the total number of trees planted in the past 12 months. We hypothesize that this is effect is due to the cumulative positive but negligible effects of the treatment on the number of trees planted in the past 12 months in each tree category, but note that we do not have enough statistical power to prove this theory. We also note that the effect of the treatment on whether the farmers planted timber trees and on the number of timber trees planted in the past 12 months is diminished by the fact that a higher share of farmers across both treatment arms planted timber trees at midline compared to baseline. This is especially noticeable in the share of farmer planting cypress trees and on the number of cypress trees planted.

Third, we find a positive and statistically significant treatment effect on the number of grevillea trees as a percentage of timber trees planted in the past 12 months. This effect is also observed in the number of grevillea trees planted as a percentage of the total trees planted in the 12 months preceding the midline. These effects are corroborated when we look at non-grevillea timber trees planted in the past 12 months as a percentage of (timber) trees planted, where we find a negative and statistically significant treatment effect.

All of the treatment effects are consistent and magnified when we look at the local average treatment effect. Farmers that partially or fully complied with the treatment protocol have a higher average treatment effect on the outcomes of interest compared to the analysis considering all treatment group farmers.

The treatment effects tend to be higher in Uasin Gishu compared to Kericho and for farmers that already owned grevillea trees at baseline. The only effect that was higher for farmers that did not own grevillea at baseline is on the total number of trees planted in the past twelve months.

We find relatively high survival rates for tree seedlings that were planted on the ground for the treatment farmers. Around two-thirds of the seeds planted from the tree kits had survived at the time of the midline data collection.

In terms of planting best practices knowledge that were asked at both baseline and midline, the treatment only had an effect on whether the farmers know which soil to use to plant tree seeds. The lack of effect on knowledge of other practices might be explained by the low attendance rate to the tree trainings. These trainings might not have been offered at all in some places according to 1AF.

However, looking at the difference in knowledge at midline between farmers that reported not attending to the training and those who reported attending, we find that the latter group tends to have better knowledge of tree planting best practices. Including knowledge questions only asked at midline, we find that farmers that reported attending the training have a marginally higher grevillea planting knowledge index. These results are statistically significant at the 5% level.

The overall attrition rate is low (6.6%), and there is no statistically significant difference in attrition rates between treatment arms. Attrition is higher in Kericho compared to Uasin Gishu. We cannot rule out that the underlying characteristics of farmers that did not participate in the midline are different of those that did because of low statistical power.

We find that a high proportion of treatment farmers (93.4%) and control farmer (97.8%) are partially compliant. However, full compliance rates are low given that a high proportion of treatment farmers did not attend the tree training. This is contextualized in the corresponding sections. We find that living in Uasin Gishu, having planted grevillea trees in the 12 months before baseline, and participating in a higher number of income-generating activities in the 12 months before the baseline are positively correlated with being fully compliant.

Appendices

Appendix 1: Comparison of baseline characteristics for farmers that planted tree kit seeds

Table 17: Baseline characteristics of treatment farmers that planted grevillea seeds

Variable	No Seedlings Planted (N=389)	Seedlings Planted (N=477)	P-Value
Respondent Gender	56%	61%	0.08*
Respondent Age	44	44	0.69
Completed Primary	71%	68%	0.29
Household Members	5.2	5.0	0.11
Number of Income Activities	2.4	2.5	0.10*
Grevillea Ownership (Y/N)	48%	46%	0.55
50 Timber Trees or More (Y/N)	52%	57%	0.08*
100 Trees or More (Y/N)	45%	51%	0.04**

Appendix 2: Regressions with different specification models

Table 18: Impact of the program on grevillea tree numbers with different specification

Dependent variable	Model	Coefficient [Std. Error]	P-Value	R^2
Number of grevillea trees planted in the last 12 months	Simple difference in differences	7.4904*** [1.3634]	0.000	0.0130
	Log difference in differences[6]	.421 *** [.0471]	0.000	0.0441
	Log difference in differences with controls	.4237*** [.0474]	0.000	0.0362
	Simple lag	6.7082*** [1.0578]	0.000	0.4388
	Lag with controls	7.8175*** [1.3865]	0.000	0.0277
No. of observations			**1,728**	

[6] The coefficients for log regressions can be interpreted using the following calculation: (exp(x) – 1) * 100. For every one-unit increase in the independent variable, our dependent variable increases by about x%. In this case, the treatment is associated with a 52.3% increase in the number of grevillea planted in the past 12 months.

Appendix 3: LATE on tree assets

Table 19: LATE on tree assets

Dependent variable	Model	Coefficient [Std. Error]	P-Value	R^2
Planted timber trees in the past 12 months	IV Regression for Partial Compliers	.0210 [.0368]	0.569	0.0015
	IV Regression for Full Compliers	.0256 [.0450]	0.569	0.0007
Planted fruit trees in the past 12 months	IV Regression for Partial Compliers	.0341 [.0346]	0.325	0.0004
	IV Regression for Full Compliers	.0418 [.0423]	0.325	0.0013
Planted other trees in the past 12 months	IV Regression for Partial Compliers	.0292 [.0271]	0.282	0.0001
	IV Regression for Full Compliers	.0357 [.0331]	0.282	0.0000
Planted any trees in the past 12 months	IV Regression for Partial Compliers	.0044 [.0345]	0.897	0.0000
	IV Regression for Full Compliers	.0055 [.0423]	0.897	0.0001
Number of timber trees planted in the past 12 months	IV Regression for Partial Compliers	5.3095 [8.8202]	0.548	0.0015
	IV Regression for Full Compliers	6.4869 [10.7849]	0.548	0.0000
Number of fruit trees planted in the past 12 months	IV Regression for Partial Compliers	1.4361 [3.3495]	0.669	0.0000
	IV Regression for Full Compliers	4.0955 [1.7339]	0.669	0.0000
Number of other trees planted in the past 12 months	IV Regression for Partial Compliers	3.5271 [2.3976]	0.143	0.0007
	IV Regression for Full Compliers	4.3093 [2.9320]	0.143	0.0002
Number of trees planted in the past 12 months	IV Regression for Partial Compliers	19.3346* [10.7892]	0.074	0.0031
	IV Regression for Full Compliers	23.6386 [13.2211]	0.074	0.0000
Grevillea trees as % of timber trees planted in the past 12 months	IV Regression for Partial Compliers	.1339*** [.0176]	0.000	0.0206
	IV Regression for Full Compliers	.1643*** [.0218]	0.000	0.0150
Grevillea trees as % of all trees planted in the past 12 months	IV Regression for Partial Compliers	.1079*** [.0141]	0.000	0.0203
	IV Regression for Full Compliers	.1323*** [.0175]	0.000	0.0109
Non-grevillea timber trees as % of timber trees planted 12 months	IV Regression for Partial Compliers	-.1339 [.01763]	0.000	0.0206
	IV Regression for Full Compliers	-.1643 [.0218]	0.000	0.0150
Non-grevillea timber trees as % of all trees planted 12 months	IV Regression for Partial Compliers	-.1135 [0308]	0.000	0.0000
	IV Regression for Full Compliers	-.1392 [.0275]	0.000	0.0000

Appendix 4: Heterogenous Treatment Effects

Table 20: Impact of the program by timber tree ownership at baseline

Dependent variable	>50 timber trees at baseline	Coefficient [Std. Error]	P-Value	R_2	No. of observations
Planted grevillea trees in the past 12 months (Yes/No)	No	.1335*** [.0275]	0.000	0.0424	785
	Yes	.1761*** [.0243]	0.000	0.0684	945
Number of grevillea trees planted in the past 12 months	No	6.2807*** [1.7519]	0.000	0.0222	785
	Yes	8.9627*** [2.3785]	0.000	0.0167	945
Planted timber trees in the past 12 months (Yes/No)	No	-.0451 [.0421]	0.286	0.0234	785
	Yes	.0653* [.0362]	0.073	0.0222	945
Number of trees planted in the past 12 months	No	15.7208** [7.9099]	0.048	0.0200	785
	Yes	12.344 [12.672]	0.331	0.0106	945
Grevillea trees as % of timber trees planted in the past 12 months	No	.0801*** [.0197]	0.000	0.0338	785
	Yes	.1103*** [.0175]	0.000	0.0530	945
Grevillea trees as % of all trees planted in the past 12 months	No	.0598 [.0153]	0.000	0.0272	785
	Yes	.0912 [.0142]	0.000	0.0514	945

Table 21: Impact of the program by grevillea ownership at baseline

Dependent variable	Grevillea at baseline	Coefficient [Std. Error]	P-Value	R_2	No. of observations
Planted grevillea trees in the past 12 months (Yes/No)	No	.1281*** [.0186]	0.000	0.1058	910
	Yes	.1885*** [.0335]	0.000	0.0400	820
Number of grevillea trees planted in the past 12 months	No	6.7575*** [1.6432]	0.000	0.0279	910
	Yes	8.8273*** [2.6937]	0.001	0.0116	820
Number of fruit trees planted in the past 12 months	No	5.9923* [3.3092]	0.072	0.0076	910
	Yes	-4.5758 [0.247]	0.247	0.0146	820
Number of trees planted in the past 12 months	No	26.6124** [11.3025]	0.019	0.0112	910
	Yes	-.6200 [11.6141]	0.957	0.0148	820
Grevillea trees as % of timber trees planted in the past 12 months	No	.0768*** [.0133]	0.000	0.0779	910
	Yes	.11833*** [.02398]	0.000	0.0313	820
Grevillea trees as % of all trees planted in the past 12 months	No	. 0741 [.0110]	0.000	0.0789	910
	Yes	.0800 [.0190]	0.000	0.0227	820

Appendix 5: Random Forest Regression

In addition to the difference in differences-, difference with a lag-, and instrumental variable regressions, we test the robustness of our results using a random forest model. The random forest algorithm is a machine learning technique that helps us to estimate the treatment effect and categorize the ranking of baseline characteristics that best predict a higher treatment effect. The results from this model are consistent with the findings reported above. Table 22 shows the main results of the model and the ranking of baseline characteristics associated to them.

Table 22: Impact of the program using random forest

Dependent variable	Coefficient [Std. Error]	Predictors (at baseline)
Planted grevillea trees in the past 12 months	0.1401 [.0252]	1. Total number of trees 2. Farmer age 3. HH Size 4. Self-reported land size 5. Ward of residence
Planted timber trees in the past 12 months	0.0010 [0.0337]	1. Total number of trees 2. Self-reported land size 3. Farmer age 4. HH size 5. Number of income activities
Planted any trees in the past 12 months	0.0203 [.0309]	1. Total number of trees 2. Farmer age 3. Self-reported land size 4. HH size 5. Number of income activities
Number of grevillea trees planted in the past 12 months	7.4564 [1.8346]	1. Total number of trees 2. Farmer age 3. Self-reported land size 4. HH size 5. Completed secondary education (Yes/No)
Number of timber trees planted in the past 12 months	3.6985 [7.9746]	1. Total number of trees 2. Self-reported land size 3. Farmer age 4. HH size 5. Knowledge of grevillea germination time
Number of trees planted in the past 12 months	18.8243 [9.5434]	1. Total number of trees 2. Self-reported landsize 3. Farmer age 4. HH size 5. Number of income activities

Grevillea trees as % of timber trees planted in the past 12 months	0.0737 [0.0172]	1. Total number of trees 2. Self-reported land size 3. Farmer age 4. HH size 5. Knowledge of grevillea germination time
Grevillea trees as % of all trees planted in the past 12 months	0.0678 [.01409]	1. Number of trees 2. Farmer age 3. Self-reported land size 4. HH size 5. Knowledge of grevillea germination time
Non-grevillea timber trees as % of timber trees planted in the past 12 months	-0.0739 [.01724]	1. Number of trees 2. Self-reported land size 3. Farmer age 4. HH size 5. Knowledge of grevillea germination time
Non-grevillea timber trees as % of all trees planted	-0.0846 [.02649]	1. Number of trees 2. Farmer age 3. Self-reported land size 4. HH size 5. Knowledge of grevillea germination time

Appendix 6: Knowledge Questions

Survey question	Training manual answers	Correct choices in the survey
What type of soil should you use for planting tree seeds?	A mixture of **top-soil**, **pure sand**, and **well decomposed manure or compost**. Consider using topsoil underneath an old healthy Grevillea tree. This soil has special fungi that can help the tree to grow quickly ***Do not use clay soil!***	Farmer mentions gathering soil from close to a healthy tree, Farmer mentions using a mix of top soil and sand, Farmer mentions using manure or compost, Farmer mentions using loam, Farmer mentions gathering soil from close to a healthy tree are **all correct answers**
What is the longest you should wait for grevillea tree seeds to germinate?	Up to 2 months!	At least a month but less than two months **is the correct answer**
How often should you water planted tree seeds if it does not rain?	If there is no rain, water 2 times per day every day	Once or twice a day **is the correct answer**
At what point should you move tree seedlings to sockets, or undertake socketing?	When there are 2 round leaves, that is around 2-3 weeks after planting	When there are 2 round leaves and Around 3 weeks after planting are **all correct answers**
What should you do before transferring seedlings to the sockets?	Prepare the tree sockets- involves preparing the potting mixture. Water the tree bag at least 20 minutes before removing tree seedlings.	Prepare sockets with topsoil and fertilizer, Be sure to move in the morning, Move tree bag to an area for easy transfer, Water the tree bag with the seedlings, are **all correct answers**

What can you do to maximize the chance of tree seedlings surviving to maturity?	Create a tree shed to protect the seedlings from damage by animals, and water the sockets in the absence of rain	Keep them in the shade, Surround them with branches/thorns etc to protect from animals, Use ash to protect against insects, Water them sufficiently, Plant trees during the rains, Apply tree fertilizer, Remove weeds, are **all correct answers**
After socketing the tree seedlings, how many weeks do you wait until hardening them?	6 weeks after socketing	
How many weeks should tree seedlings be hardened before planting in a field?	At least 2 months	
How far apart should trees be spaced when planting?	One meter apart, that is two bean sticks	Farmer mentions 1 meter or more is **the correct answer**
When pruning trees, what proportion of the branches on one tree is it safe to cut without causing potential harm to the tree?	Prune the bottom 25% or less of the branches on the tree	Farmer mentions 25% (a quarter) of the branches or fewer, is **the correct answer**
How many months old should a tree be before the first pruning?	The first prune should be done when the tree is 12 months old	

laterite

DATA | RESEARCH | ADVISORY

From data to policy