Assessing the Downstream Socioeconomic and Land Health Impacts of Agroforestry in Kenya

Impact Assessment Report

Preparation date: 02/08/2017

Prepared by: Karl Hughes[1], Seth Morgan[2], Kathy Baylis[2], Judith Oduol[1], Emilie Smith-Dumont[1], Tor-Gunnar Vagen[1], Mary Mutemi[1], Claire LePage[3], and Hilda Kegode[1]

[1]World Agroforestry Centre (ICRAF)
[2]University of Illinois, Urbana-Champaign, United States of America
[3]Independent Research Fellow

Contents

Abstract 1
1. Introduction 2
2. Vi Agroforestry's Agroforestry Promotion Programme and Implicit Theory of Change 4
2.1 Vi Agroforestry's Programme 4
2.2 Programme Theory of Change 6
3. Impact Evaluation (Causal Identification) Strategy and Implementation 8
3.1 Impact Study Area Selection 8
3.2 Impact Evaluation Design 9
3.3 Farmer Group Identification and Sampling 16
3.4 Data Collection and Analysis 16
3.5 Baseline and Time Invariant Respondent and Household Covariate Balance 18
4. Empirical Results 20
4.1 Connections between ICRAF/CGIAR research and Vi's programme 20
4.2 Vi Programme Exposure 22
4.3 Adoption of Promoted Agroforestry and Related Practices and Tree Germplasm 25
4.3.1 Agroforestry Practices and Tree Species 25
4.3.2 Other Sustainable Land Management Practices 31
4.4 Remote Sensing Derived Tree and Vegetation Cover Estimates 34
4.5 Remote Sensing Derived Soil Health Estimates 38
4.6 Tree Product Sales and Firewood Access 42
4.6.1 Tree Product Sales 42
4.6.2 Firewood Cash Value and Collection Time 46
4.7 Tree Fodder Use and Milk Yields 49
4.8 Household Consumption Expenditure and Asset Wealth 53
4.9 Coping Strategy Index and Food and Nutritional Security 63
4.10 Education Progression, Spending & the Economic Ladder 66
4.11 Investigating mechanisms for key greater asset accumulation among households in the Programme Area 69
5. Discussion and Policy Implications 80
References 83
Annexes 87
Annex 1: Village Matching using Spatial and Secondary Data 88
Annex 2: Baseline and Time Invariant Characteristic Comparisons 91
Annex 3: Programme Exposure Tables 95
Annex 4: Agroforestry Adoption Index 97
Annex 5: Differential Effect Estimates 101
Annex 6 : Covariate Comparison of HHs in Programme and Non-programme Areas—Female Participants Only 122

Abstract

Agroforestry is widely purported to improve the livelihoods of smallholder farmers, rehabilitate degraded landscapes, and enhance the provisioning of critical ecosystem services, including carbon sequestration. Yet, the evidence base supporting these longer-term impacts is weak. Using a quasi-experimental evaluation design based on a theory-based and mixed methods framework, we investigated both the downstream and intermediate effects of a nine year effort led by Vi Agroforestry (herein Vi), a Swedish non-governmental organization (NGO), to promote agroforestry in large sections of Bungoma and Kakamega counties in western Kenya. In particular, we compared two sets of households against various outcome measures along the causal chain: those belonging to (a) 226 pre-existing farmer groups operating in 60 targeted programme villages; and (b) 206 pre-existing farmer groups operating in 61 geospatially matched comparison (non-programme) villages. To further counter selection bias, we combined several econometric analytical methods, including two-stage least squares regression (2SLS), with difference-in-differences estimation. In addition, to triangulate key findings and interrogate impact pathways, unforeseen outcomes, and unexpected quantitative results, we carried out semi-structured in-depth interviews with a sub-sample of 40 purposively selected programme participants. We also applied process tracing to investigate the linkages between Vi's programme and previous agroforestry research carried out by the World Agroforestry Centre (ICRAF). We found these research-to-programme linkages to be strong, and that greater—albeit variable and generally modest—programme exposure and agroforestry uptake took place among the farmer groups Vi targeted. Similarly, significant, yet again variable, effects were also identified for agroforestry product income, fuelwood access, and milk yields among dairy farmers. Soil organic carbon (estimated through remote sensing) increased at a higher rate overall in the sampled farm plots of the programme villages, but, ironically, so too did soil erosion. Finally, we found limited evidence that the programme significantly bolstered food security, shock resilience, and education progression and spending. However, we identified statistically significant—although, again, modest—programme effects for our asset and consumption expenditure measures (which includes our primary outcome variable), particularly among households represented by female programme participants.

1. Introduction

Relatively few studies have rigorously assessed the longer-term impacts of agroforestry extension programmes and integrated agroforestry systems. A number have, however, examined the effectiveness of specific agroforestry practices on intermediary outcomes, such as soil fertility and crop yields (e.g. Odhiambo et al. 2001; Otsuki 2010; Sjögren et al. 2010; Sileshi et al. 2009). The results are mixed, indicating that the effectiveness of agroforestry on such outcomes is largely dependent on the specific practices introduced, the extent to which they are appropriately implemented, and their interaction with the biophysical and socioeconomic context in question. Fewer studies have attempted to examine the effects of specific agroforestry interventions or integrated systems on more downstream outcomes, such as household income and food security. Place et al. (2005) attempted this in relation to agroforestry-based soil fertility replenishment practices in western Kenya using instrumental variable estimation. However, the sample of households studied was small (n=102), and the instruments used (e.g. whether any adult in the household previously held a job) may violate the exclusion restriction[1], thereby rendering the results largely inconclusive.

This impact assessment is motivated by a dearth of evidence on the more distal impacts of integrated agroforestry programmes and systems.

Moreover, farmers typically do not follow only one agroforestry practice for one specific purpose. Rather, multiple practices are pursued simultaneously, with the intention of deriving multiple benefits, e.g. soil and crop management, fodder, fuelwood, fruits, and timber for sale. Many of these practices are further expected to positively interact (Nair 1993). Consequently, while single practice efficacy studies are clearly important, there are limitations in relying upon them alone to understand the impacts of agroforestry.

Given this dearth of evidence—coupled with its increasing policy prominence (see, for example, Buttoud et al. 2013)—further efforts to understand the downstream impacts of agroforestry are important. One possible approach would involve promoting contextually appropriate agroforestry practices in randomly selected villages for a significant number of years and then comparing households within them with those in control villages against various intermediate and downstream outcome measures. Yet, executing such a study would be challenging, particularly given the time it takes for the potential impacts of many agroforestry practices and their synergistic interactions to fully manifest, coupled with the likelihood of significant spill over effects and contagion during this time period. Identifying an appropriate instrumental variable that only affects the outcomes of interest via influencing the uptake of agroforestry would be the next best thing. However, being in the fortunate position to have good instruments is rare, and their exogenous nature is untestable, thereby making the resulting treatment effect estimates difficult to defend (Blundell and Costa Dias 2009).

Nevertheless, there are geographic areas where specific agroforestry practices have been promoted both intensively and for considerable periods of time. The impact assessment documented by this report sought to take advantage of the existence of such areas in western Kenya targeted by Vi Agroforestry (herein Vi), a Swedish non-governmental organization (NGO). In the next section, Section 2, we describe this programme, including our rationale for selecting it and its implicit theory of change.

[1] The exclusion restriction states that the instrument in question should only affect the outcome variable indirectly by influencing treatment status (Angrist and Pischke 2008).

Given that our impact assessment was carried out following the implementation of Vi's programme, i.e. it was not planned in advance, key opportunities were missed to promote 'clean' causal identification. Consequently and as is the case with most quasi-experimental designs and observational studies, we had to go to considerable lengths to overcome the inherent limitations. Section 3, in particular, outlines the multiple approaches we used to overcome both programme placement and self-selection bias, as well as our sampling, data collection, and data analysis strategies.

Section 4—Empirical Results—is the heart of this report. Here, the results for each step in the theory of change of Vi's programme are presented. This starts with a summary of a sub-study that was carried out to interrogate the causal linkages between ICRAF's research efforts in the 1990s and early 2000s on the one hand and the agroforestry practices and tree/shrub germplasm Vi promoted under its programme on the other. We then assess the extent to which the targeted smallholder farmers participated in this programme and, in turn, took up the promoted practices and germplasm. An examination of the **intermediary outcomes**—ranging from remote sensing derived measures of tree cover and soil health to firewood access, tree fodder use, milk yields, and revenue from the sales of agroforestry products—and **final outcomes and impacts**—ranging from consumption expenditure and asset accumulation to shock reliance, food security, and education progression spending—along the causal chain then follows. Section 5 concludes with a summary discussion and policy implications.

2. Vi Agroforestry's Agroforestry Promotion Programme and Implicit Theory of Change

2.1 Vi Agroforestry's Programme

General background:

Vi is a Swedish NGO founded in 1983, which operates in four African countries—Kenya, Uganda, Tanzania, and Rwanda. Over the years, Vi has promoted several interrelated, complementary agroforestry practices. This includes woody perennials for (a) domestically consumable and marketable products, e.g. timber, fuelwood, and fruits; (b) natural resource management (NRM) enhancement, e.g. soil fertility improvement, soil erosion control, and increased water infiltration; and (c) livestock fodder. Its current programme model focuses on promoting the above among pre-existing smallholder farmer groups, coupled with other sustainable agricultural land management (SALM) practices, e.g. composting, crop rotation, and mulching. This is combined with complementary farmer group capacity development activities, e.g. leadership training and the promotion of group savings and lending.

Vi's programme, as implemented in the impact study area, is primarily focused on the promotion of agroforestry, together with other complementary sustainable land management practices.

Vi received extensive training and support from the World Agroforestry Centre (ICRAF) throughout the 1990s and early 2000s, including a series of seminars, which incoming Vi extension staff and managers received from the mid-1990s up until 2005. Training from ICRAF during this period, for example, introduced Vi to an agroforestry practice known as alley-cropping, which involves the planting of leguminous shrubs in rows in between annual crops, with the intention of providing additional fertility through nitrogen fixation and the incorporation of leafy biomass (Douthwaite et al. 2002). ICRAF also introduced Vi staff to the use of perennial legumes, including *Calliandra callothyrsus* and *Sesbania sesban*, as fodder and improved fallow species (LePage Morgan 2017). (See Subsection 4.1 for further details on the linkages between ICRAF's research and Vi's agroforestry promotional efforts.)

The agroforestry techniques Vi primarily promotes can be described as variants of three separate planting patterns: alley-cropping (described above); boundary planting; and tree planting along and to strengthen erosion control structures. Boundary planting is a common practice throughout both the targeted and non-targeted parts of the impact study area, but Vi encourages farmers to intensify this practice and add more short-term leguminous shrubs in the spaces in between the traditionally planted long-term timber species, thereby creating a multi-story boundary planting system (Wachiye 2008; LePage Morgan 2017). Additionally, farmers are trained to develop similar multi-story perennial systems along intra-plot erosion control structures, which include simple grass strips, trash lines consisting of crop residue, small contour bunds, trenches, and terraces. This practice of planting trees for erosion control is described by one promoter as a "slow terrace," and is reported as less labour-intensive than hand-built earth terraces (LePage Morgan 2017).

Vi's self-described hallmark is the incorporation of leguminous perennials, especially *Sesbania sesban* and *Calliandra callothursus*, into the farming system. One senior and long term Vi staff member reported that: "If you see *Sesbania* you can guarantee that Vi had been there" (E. Wachiye, personal communication, March 2016). These trees

are fast-growing and nitrogen-fixing, and they are used for firewood, fodder, and 'green manure' through the incorporation of the leafy biomass into the soil for to enhance soil fertility and structure.

In practice, Vi's programme participants tend to pick and choose from among the promoted species and planting arrangements, while adapting them to their specific needs and circumstances. It is further worth noting that the implementation of Vi's programme is not particularly standardized. Its extension staff, in particular, are given latitude to assess the topics needed by their assigned farmer groups. Nevertheless, every group is expected to learn about the advantages of agroforestry, and Vi's activity calendar is coordinated around its twice-yearly tree seed distribution, which corresponds with the arrival of the two rainy seasons in its operational area. These tree seeds are distributed free of charge to all its member groups. These include seeds for direct-seeding in fields before the long rains as maize is planted, as well as seeds for raising in small-scale tree nurseries before the short rains when secondary bean crops and vegetables are planted.

It is also worth noting the changing nature of Vi's extension approach over the years. In the early 1990s, a large contingent of 400 extension officers was recruited and deployed in the organization's 'areas of concentration' at the village level to promote more technical and context-specific agroforestry practices among farmers. This included training farmers on how to identify tree seeds and set up and manage their own nurseries. This continued until 2004 where the number of extension officers was significantly reduced and farmers were supported in groups, rather than individually.

The farmer-led groups through which Vi delivers its training and seed provision are common throughout western Kenya. They tend to have 10-15 members, which are often organized as women, youth, farmer, or religiously focused groups. However, in practice, these distinctions are not binding. Nearly all groups contain a mix of ages and gender, but women tend to predominate, and most group members are older and slightly better off than the general population.

Specific Projects Implemented in the Impact Study Area:

The two projects implemented in the programme area are similar but have their own unique areas of focus—carbon sequestration on the one hand and farmer organization capacity development on the other.

Vi's agroforestry promotion efforts have taken place in the impact study area since 2008 through the implementation of two projects: the Kenya Agricultural Carbon Project (KACP) and the Farmer Organizations and Agroforestry (FOA) project. (See Figure 3.2.1.) These two projects have their own field staff and funding structures but share similar approaches for promoting agroforestry and other complementary SALM practices.

The distinguishing feature of KACP, however, is that it explicitly emphasises the carbon sequestration function of the promoted agroforestry and other SALM practices. Here, tree planting and follow-up care is incentivized by a small payment (equivalent to approximately USD $3.00 per person per year, on average) disbursed to the farmer groups upon confirmation that trees have been planted and cared for on their farms.

FOA, on the other hand, stresses the capacity building of farmer organizations, as a complement to the provision of tree seeds and land management training. Vi staff

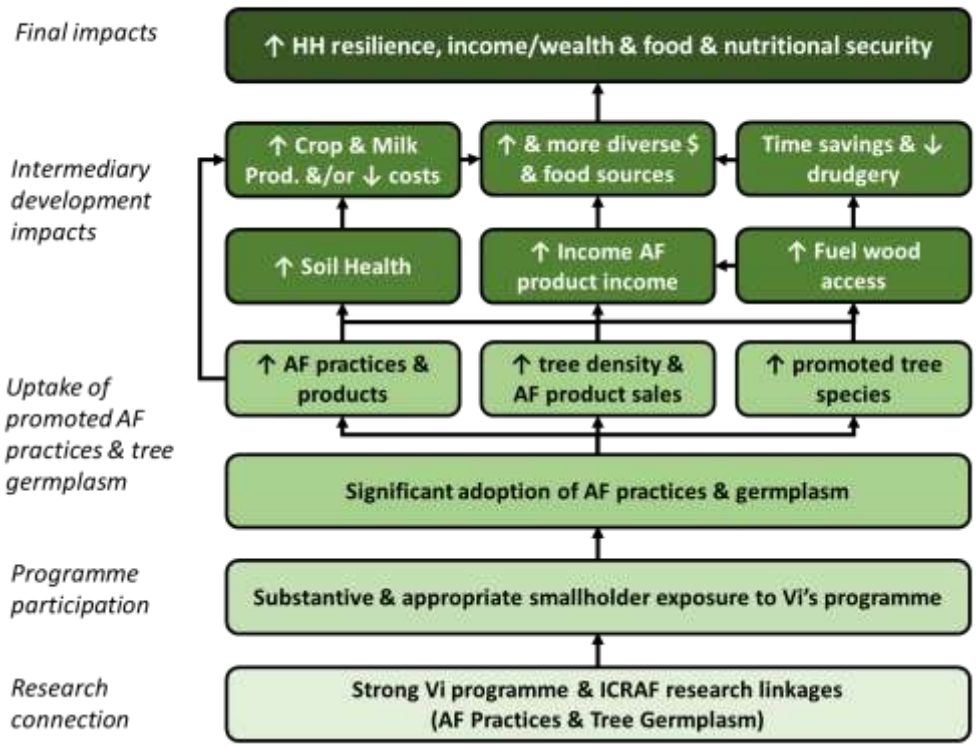
describe the extension training provided under FOA as essentially the same as that provided under KACP, but with a few key differences: FOA does not provide carbon payments, nor monitor tree planting with the same degree of rigour. Moreover, since FOA is focused on empowering farmer organizations, as decision was made in 2014 to hand over its activities in the centre of the impact study area (i.e. the Kimilili/Ndivisi Village Sampling Zone [VSZ]) to partnering Savings and Credit Cooperatives (SACCOs). This means that Vi training and other capacity development activities for farmer groups have been carried out through SACCOs for the latter three years of the study period in one of the four VSZs associated with our study, i.e. from 2014 to 2016. Despite the above differences, the training and seed distribution regimen over the two projects has largely been the same, and Vi's expectation is that all their participants will become successful adopters of agroforestry.

2.2 Programme Theory of Change

To help structure the impact assessment, we constructed a basic theory of change (Figure 2.2.1) for how Vi's programme is expected to have worked to generate its expected longer-term impacts, including the link back to ICRAF's agroforestry research efforts.

Given that we sought to carry out a theory based impact assessment, a key initial step was to (re) construct a theory of change for Vi's programme.

FIGURE 2.2.1: Theory of Change Framework for Vi's Programme



Given that a key objective of our impact assessment is to assess the linkages between ICRAF's research efforts and the agroforestry practices and tree/shrub germplasm, the starting point of the theory of change is an underlying assumption (or precondition) that these linkages existed and were strong. We further need to assume that both meaningful and appropriate participation in Vi's programme by the targeted farmer groups took place, another essential precondition for the manifestation of the

proceeding expected outcomes and impacts. This is especially important given the complex and multi-faceted nature of agroforestry. With such substantive programme participation, a high uptake of the specific agroforestry practices and tree/shrub germplasm promoted among the participating smallholders is then expected to have followed. As per our Agroforestry Index, introduced in Subsection 4.3, this would be reflected by: (a) the extent of practice uptake, i.e. of the promoted practices and use of the resulting agroforestry products; (b) intensity of practice, as reflected by increased on-farm tree cover and the selling of agroforestry products; and (c) the presence the 'signature' tree and shrub species that Vi promoted, e.g. via its bi-annual seed distribution activities.

With such adoption, we would expect to see a number of intermediary impacts materialize. One is improved soil health, which would be expected to, in turn, increase crop production, or—at the very least—reduce input costs and, therefore, increase returns. And with the increased use and availability of tree/shrub fodder, increases in milk production and/or returns would also be expected among dairy farmers. Moreover, increases in revenue from other agroforestry products, such as timber, firewood, and fruit, is further expected to have increased and diversified income and food sources. Given their traditional labour activities, benefits specific for women would additionally be expected, due to the increased availability of on farm firewood. The above intermediary impacts were then expected to have interacted together to bolster household income, food and nutritional security, and resilience to shocks.

With this background of Vi's programme, we will now describe how we sought to evaluate its outcomes and impacts in general and the extent to which the above theory of change unfolded as expected in particular.

3. Impact Evaluation (Causal Identification) Strategy and Implementation

Given that the rollout of Vi's programme was not informed by an appropriate, prospective impact evaluation design, we went to considerable lengths to both devise and implement several strategies to enable an evaluation of its outcomes and impacts in the most credible way possible. The purpose of this section is to describe these strategies, including the rationale for pursuing each and how we implemented them.

3.1 Impact Study Area Selection

Our quest to assess the impacts of an agroforestry promotional effort in Kenya began even before we engaged Vi in this exercise. We started first by exploring locations where agroforestry had been substantially promoted in various areas of the country where a credible quasi-experimental impact assessment design could be pursued. Our scoping exercise revealed that, while there had been a number of intensive, long term efforts to promote agroforestry in Kenya, this was often done in combination with other interventions, thereby making it difficult to evaluate its specific impacts. The only organization found to have had a sustained and near exclusive focus on agroforestry promotion was Vi, hence the genesis of this particular impact assessment.

However, before pinning down a specific evaluative approach (i.e. a causal identification strategy), we first sought to understand the history of Vi's programme and how it was rolled out, both spatially and temporally. There was some hope that this *could have* led to the identification of one or more plausible instrumental variables, i.e. one or more quasi-random factors that influenced exposure to and/or uptake of the promoted agroforestry practices and tree/shrub species but otherwise unrelated to our outcomes of interest. Unfortunately, one of the initial instruments explored—something that would have exploited features of the historical land redistribution process that happened in Vi's home-base, Trans Nzoia County—turned out to be unviable. Our attention, therefore, shifted to identifying areas where Vi had been working for a number of years and matching these up with comparable areas where it had not, as well as where there had not been any other significant efforts to promote agroforestry.

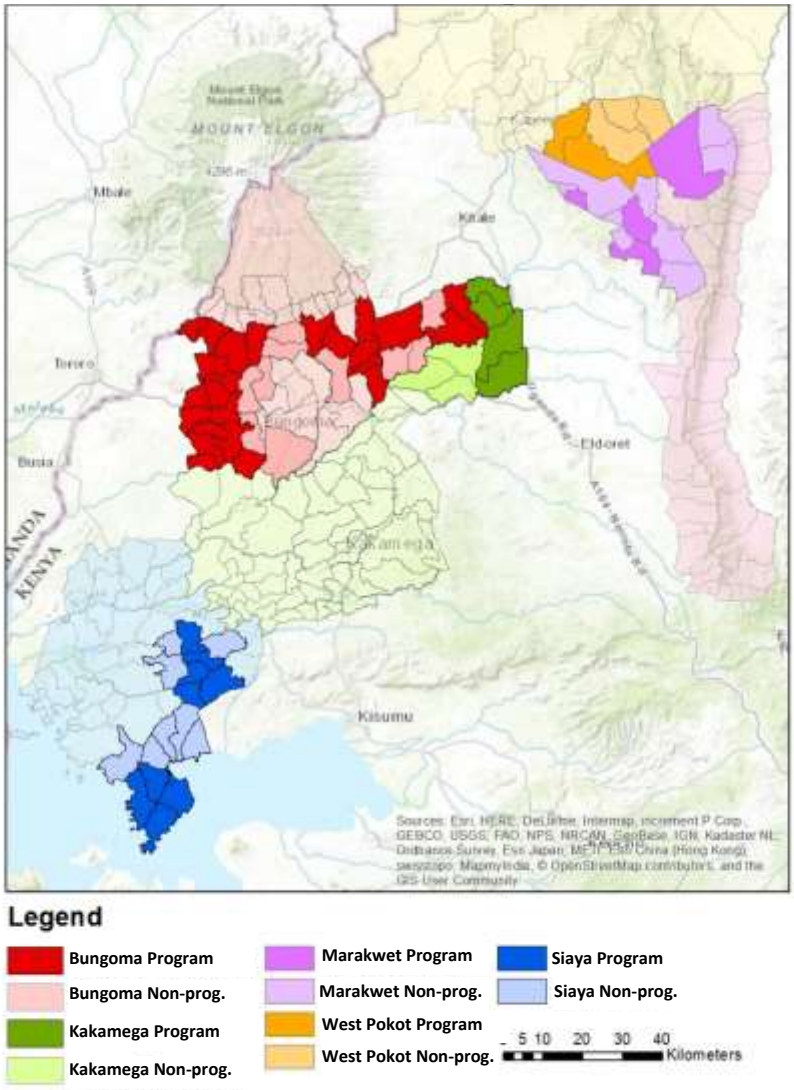
Our first strategy involved identifying similar geographic areas not targeted by Vi or any other substantial agroforestry promotional effort.

This culminated in a scoping mission that we carried out in March 2016—together with Vi staff—to further narrow in on suitable intervention and comparison areas to serve as the study's focus. Following initial discussions, a decision was made to visit three potential programme areas—(1) West Pokot and Marakwet counties; (2) Bungoma and Kakamega counties; and (3) Siaya County (see Figure 3.1.1). These areas were shortlisted because Vi had been operational within them for a number of years and where the potential of identifying plausible comparison areas existed.

West Pokot and Marakwet counties were ruled out, given that the uptake of agroforestry was reported by Vi staff to be low, which was corroborated by our observations in the field. Explanations for this by Vi staff included the failure of fodder tree species to thrive in this highland agro-ecological setting and the presence of forest areas decreasing the demand for timber grown on farm. Siaya County was also disfavoured due, again, to reportedly low rates of adoption, as well as difficulty in identifying credible comparison areas with similar agro-ecological characteristics.

Figure 3.3.1: Potential Study Impact Areas Scoped

The counties of Bungoma and Kakamega were selected to serve as the impact study area, given their reportedly higher levels of agroforestry adoption and the presence of comparable areas where substantive agroforestry promotion had not taken place.



We found the greatest potential in Bungoma and Kakamega counties. After visiting a number of sites and conducting informal interviews with farmers, we concluded that this area had high potential for the impact assessment, given the (reportedly) high rates of agroforestry adoption and agro-ecological comparability across programme and potential comparison areas. In addition, Vi had not operated in the area prior to 2008, i.e. our study's baseline period, and no other organization had substantively promoted agroforestry in the area.

3.2 Impact Evaluation Design

Given the study's non-experimental nature, coupled by the fact that an appropriate baseline survey was never undertaken, identifying a geographical area with perceived high rates of programme uptake and potential areas for comparison purposes was only an initial first step in ensuring a strong quasi-experimental impact assessment design. In particular, we had to undertake a number of additional measures to counter both programme placement and self-selection bias (White 2010), as well as other

internal validity threats, i.e. programme spill-over effects and contagion. These included:

1. *Village matching based on selected geospatial and demographic variables.*
   During the initial stages of its programme, Vi targeted specific geographic areas (i.e. Locations) and then pre-existing and active farmer groups within them. Almost all of these groups operated in specific villages or clusters of neighbouring villages. To counter potential programme placement bias in particular, we undertook the following steps:

We mitigated programme placement bias by matching programme and comparison villages based on key geospatial and demographic variables.

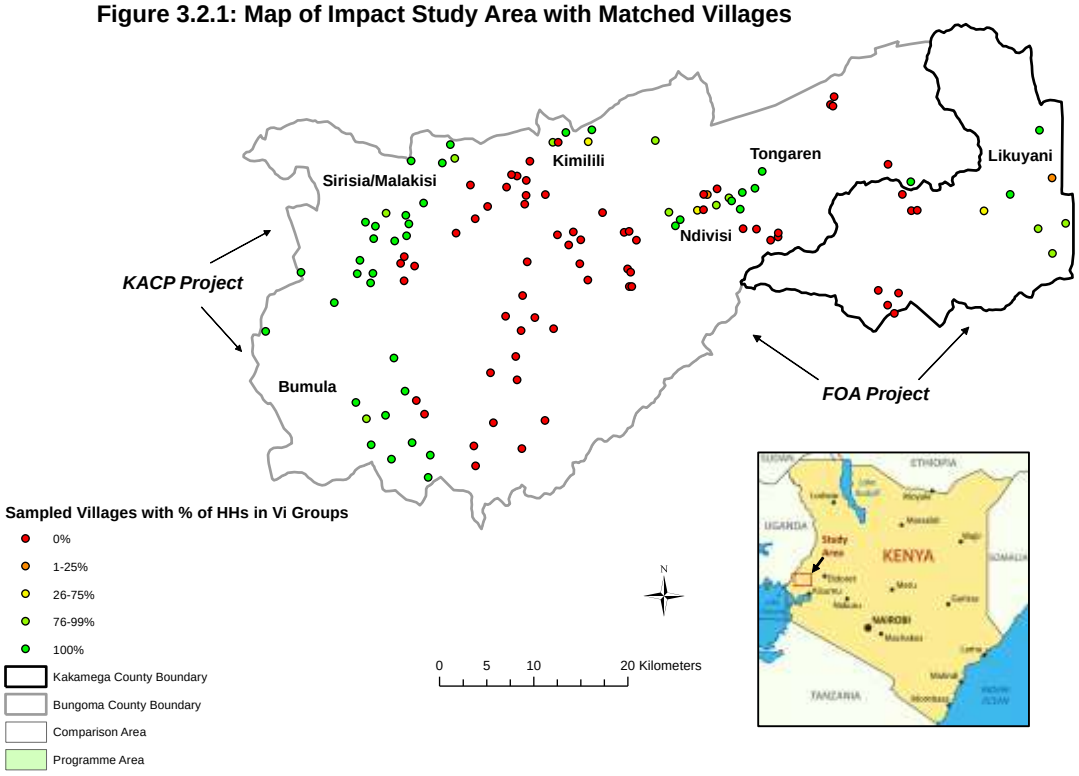
   a) We identified specific sublocations (the smallest administrative unit above the village in Kenya) where Vi had operated since the baseline period. We then worked with local informants to purposively match these with potential comparison sublocations based on their similarity in terms of perceived wealth status and agro-ecological characteristics.
   b) A scoping survey was administered in all the villages within the purposively matched programme and comparison sublocations. This included capturing the villages' central geographical positioning system (GPS) coordinates and verification of the existence of active farmer groups that had been operational since the initial years of the Vi's programme.
   c) The GPS coordinates were used to obtain village specific estimates on key geospatial and demographic variables from secondary data—i.e. population density, baseline soil conditions and tree cover, elevation, rainfall, and distance from major road networks (as a proxy for market access).
   d) Propensity score matching (PSM) (Rosenbaum and Rubin 1983) was then used to identify a set of programme and comparison villages balanced across these variables. In particular, we aimed to identify 30 comparable programme and comparison villages in equal numbers within each primary Vi supervision area, hereafter referred to as Village Sampling Zone (VSZ). Using the *psmatch2* programme (Leuven and Sianesi 2003) in Stata, one-to-one calliper matching was first implemented at the VSZ level. Here, the callipers employed were incrementally reduced, thereby iteratively discarding poorly matched programme and non-programme villages until the sample of 30 well-matched villages (15 programme villages and 15 comparison villages) per VSZ was reached. In the end, the initial sample of 336 villages (194 programme and 142 comparison villages) was reduced to 121 (60 programme and 61 non-programme), with one additional village added given that one of the matched comparison villages turned out to actually be two distinct villages. Figure 2.2.1 presents the locations of the final set of matched villages, while Annex 1 describes this village matching exercise in greater detail. It is worth noting that the above village matching exercise served as the basis of a MSc student's Masters thesis (Morgan 2017).

2. *Random selection of respondents from all farmer groups that existed in both the intervention and comparison villages since the baseline period.*
   We pursued this strategy to explicitly counter self-selection bias. As mentioned above, Vi initially targeted pre-existing farmer groups, as opposed to mobilizing new groups. There was likely significant self-selection at play in the formation of these groups. This would potentially bias a comparison of these groups with random samples of farmers in the comparison villages. Consequently, only households belonging to active farmers groups existing since the baseline period

One of our strategies for mitigating self-selection bias involved capturing data from *all* pre-existing farmer groups in both the programme and comparison areas.

were sampled in the comparison villages. The assumption here is that Vi would have offered these groups the opportunity to participate in its programme had there been sufficient resources to expand into these particular villages. Moreover, in the event, not all the pre-existing farmer groups that were targeted in the programme villages took up the opportunity to participate in Vi's programme. Thus, those groups that did are actually a sub-set of the targeted farmer groups, which may be unique in terms of both their observable and unobservable characteristics relevant to our outcomes of interest. Bias could, therefore, have crept in if we had directly compared them with the sample of pre-existing farmer groups in the comparison villages. We therefore sampled households from *all* farmer groups that existed in *both* the intervention and comparison villages at baseline, while seeking to identify two types of causal effect estimates described under point 5 below.

Figure 3.2.1: Map of Impact Study Area with Matched Villages



3. *Reconstruction of baseline data for difference-in-differences estimation*

Baseline data—through respondent recall and analysis of 2007/08 satellite imagery—enabled us to control for pre-existing baseline differences between the programme and comparison areas.

A key limitation of our study is that no suitable baseline survey was undertaken. Consequently, differences in outcome found between the farmer groups in the matched programme and comparison villages may simply reflect their pre-existing baseline differences. The absence such data also prevents difference-in-differences estimation, an appropriate causal identification strategy when it is reasonable to assume that the rate of outcome change experienced between the two groups would have been the same in the counterfactual state (i.e. the parallel trend assumption) (Abadie 2005). However, during data collection, recalled baseline data were obtained from the respondents, particularly for items we assume recall bias would be minimal. This included ownership of assets, other wealth indicators (e.g. housing characteristics), household livelihood pursuits, and tree planting and land management practices. We took advantage of the fact that

a significant historical event took place one year prior to the baseline period, i.e. Kenya's post-election violence, and used this as a historical marker. The respondents were specifically requested to recall conditions prior to the nation-wide events that took place in December 2007 to January 2008. The geographical area where the survey was carried out did not witness significant violence, as had taken place in other parts of Kenya, so we thought it was appropriate to take advantage of this particular historical marker.

The recalled baseline data were therefore used to generate difference-in-differences estimates for several important outcome measures, e.g. changes in asset wealth. For obvious reasons, consumption expenditure data were only captured for the endline period. However, the rich basket of 91 assets and other wealth indicators obtained for both periods was used, together with both the asset and consumption expenditure endline data, to derive a consumption weighted asset measure (see Subsection 4.8). In addition, we compared the groups in relation to four farm plot-level differenced biophysical measures—tree cover, fractional vegetation cover[2], soil organic carbon, and soil erosion. As further elaborated below, these measures are based on the analysis of 30x30m satellite imagery using predictive models derived from ground-truthed data associated with ICRAF's Land Degradation Surveillance Framework (LDSF) field sites.

4. *Econometric modelling to control for other observable differences between the two groups.* The Achilles heel of the difference-in-differences estimator is the parallel trend assumption; that is, such estimates are biased if the outcome of interest would have changed at a different rate among either the intervention or comparison group in the counterfactual state. This could be through either groups' differential exposure to external events or shocks and/or due to the influence of their unique characteristics, whether observable or unobservable. We therefore sought to strengthen the difference-in-differences identification strategy by combining it with modelling methods that control for observable differences (such household education and baseline wealth status) that may have caused a violation of the parallel trend assumption. This would not, however, have addressed bias resulting from differential external event/shock experience and/or the influence of unobservable characteristics (e.g. business acumen) in affecting the evolution of the outcomes of interest over time.

We used several complementary modelling approaches to control for the 'residual' baseline and time invariant differences found between respondents and households of the programme and comparison areas.

Partly as a robustness check and partly as a means for better understanding the nature of the programme's estimated effects, we implemented the following modelling approaches:

- **Ordinary Least Squares Regression (OLS)**, while controlling for baseline and time invariant household and respondent characteristics (covariates) correlated with our programme area dummy variable (p<0.1).
- **Inverse Probability Weighted Regression Adjustment (IPWRA)** or doubly robust inverse probability weighted regression, given that it offers protection against misspecification of either the participation or outcome model (but not both) (Funk et al. 2011). Here, the outcome models included both these covariates of the OLS models and covariates correlated with outcome (also p<0.1) derived through stepwise regression.

[2] This is a relative index that indicates the extent to which the sampled fields are covered by green vegetation.

- **Robust regression** to mitigate the influence of extreme values in the distributions.
- **One-to-one nearest neighbour matching** as a non-parametric programme effect estimation strategy.
- **Quantile regression** to generate median, rather than average, programme effect estimates.

Given the village matching exercise (Point 1 above) was first implemented at the VSZ level and to minimize our results being influenced by general VSZ specific differences, all our linear models included VSZ dummies as fixed effects and the nearest neighbour matching imposed exact matching within each VSZ. Moreover, for all OLS and 2SLS models, standard errors were clustered at the farmer group level, given Vi's targeting approach of engaging pre-existing farmer groups.

5. *Use of Intention-to-treat (ITT) and local average treatment effect (LATE) estimation.*
   Given what is stated under Point 2 above, we generated two types of treatment effect estimates: ITT effect estimates and LATE estimates. The former were derived by comparing all sampled households in the villages targeted by Vi with those in the comparison villages, using the above estimation techniques, regardless of whether they happen to belong to a Vi group. However, given that the sample of households from the programme area includes a significant number (~25%) that are not members of Vi groups and, hence, did not directly participate in Vi's programme, the ITT estimates likely underestimate the effects of such participation.

   One possible method of obtaining a more refined estimation of the impacts of Vi's programme on those households that actually participated would have involved comparing households that belonged to the Vi affiliated groups with those households in the comparison villages that are observationally similar (statistically speaking). This would enable an estimation of the average treatment effect on the treated (ATT). However, such treatment effect estimates would rely strongly on the 'selection on observables' assumption (Handa and Maluccio 2010) and undermine the bias mitigation strategies presented under Point 1 and Point 2 above. In other words, it would fail to rule out the possibility of there being non-programme related unobservable differences between the two groups driving any identified programme effects.

We used an econometric modelling method know as 2-stage least squares regression (2SLS) to estimate programme effects on the Vi farmer group members in particular.

   Alternatively, given that the programme and non-programme villages were matched fairly successfully on key geophysical and demographic characteristics, we can assume that they are as good as randomly assigned (conditional on controlling for 'residual' baseline and time invariant differences between the two groups). That is, we can assume that the village matching process largely eliminated programme placement bias, leaving self-selection bias as our primary concern. We can further assume that the opportunity provided by Vi to the pre-existing farmer groups in the programme villages to participate in its programme made it more likely and never less likely for them to have actually participated (i.e. the monotonicity assumption). With these two assumptions, we applied two-stage least squares regression (2SLS) to derive LATE estimates (Imbens 2010). Given that there are no households in the non-programme areas that are members of Vi

groups, these effect estimates essentially pertain to those households that were members of the farmer groups that actually participated in Vi's programme.

6. *Use of mixed methods to interrogate mechanisms*
   Given that we adopted a theory-based approach, quantitative data were captured through the household survey on various intermediary measures along the causal pathway towards the programme's expected effects on consumption expenditure, food security, and resilience, following the theory of change presented in Subsection 2.2. This enabled us to assess to the extent to which changes associated with this theory of change unfolded as expected. We further complemented this with statistical mediation analysis (MacKinnon 2008) using Stata's *sem* (structural equation modelling) command, the results of which presented in Subsection 4.11. This was to assess the extent to which several hypothesised mechanisms for how particular downstream programme effect estimates came about are consistent with the variation in the data. However, it is important to emphasise that SEM, as a tool for undertaking mediation analysis, does not prove that the variables are causally related in the way they are specified in the particular model and certainly does not identify the direction of the causal relationship. Nevertheless, significant mediation effect estimates can increase confidence in the validity of the hypothesized mechanism. By extension, the converse can weaken such confidence (Hughes 2012).

   Our impact assessment also included a substantive qualitative component, which took the form of two separate sub-studies. The primary objectives of the first—carried out by a team of two qualitative researchers—was to (1) analyse variation in agroforestry adoption intensity between female and male Vi group members and across the two main project areas (KACP and FOA); and (2) explore the mechanisms through which different components of Vi's programme may have contributed to livelihood improvements. This involved carrying out semi-structured interviews with a sub-sample of 40 purposively selected Vi group members, stratified by gender and location with 20 farmers in four villages under KACP (Bumula and Sirisia constituencies) and 20 in four villages under the FOA project (Likuyani and Kimilili constituencies). Here, we used a structured questionnaire combined with ranking exercises, record sheets for trees and products, farm sketches, and in depth interviews for formal local knowledge acquisition using the Agro-ecological Knowledge Toolkit (AKT5). Two formative qualitative visits were also carried out to help both shape this work and also inform the design of our quantitative survey instrument.

Two qualitative 'sub-studies' were also carried out to interrogate variation in agroforestry adoption, the mechanisms through which such adoption could lead to the expected benefits, and the links back to ICRAF research.

   The methodology of process tracing was further used to evaluate the causal linkages between ICRAF's research on the one hand and the practices and germplasm promoted by Vi in the study area on the other. It sought to answer two specific questions:

   (1) What parts of Vi's programme were significantly informed by ICRAF's research?
   (2) How was ICRAF's research transmitted to Vi?

   In process tracing, competing causal hypotheses are subjected to a series of tests, with the goal of affirming or disaffirming each (Collier 2011). The following presents a summary of each, with specific examples relevant to this particular

study:

- A **"straw-in-the-wind"** test is the weakest process tracing test. It is neither necessary nor sufficient to confirm or reject a hypothesis, but can show which way the 'wind is blowing'. For example, a Vi training manual, created five years after the start of its programme, references an ICRAF publication that came out four years after the it started. Here, there is the suggestion that ICRAF research may have been relevant to the programme's work from the start, but there is no confirmatory power.

- A **"hoop"** test is necessary but insufficient to confirm a hypothesis. It presents, as its name suggests, a hoop that a hypothesis must jump through in order to be confirmed. Failing a hoop test eliminates a hypothesis; passing it keeps the hypothesis under consideration, and somewhat weakens rival hypotheses (Collier 2011). For example, ICRAF's research may have made a measurable contribution towards developing an agroforestry practice and trained Vi on its application; this is necessary, but insufficient, to show that ICRAF research influenced Vi's training programme, i.e. we would also need to know that Vi actually promoted it among smallholder farmers.

- A **"smoking gun"** test is sufficient but not necessary to confirm a hypothesis. Imagine a crime scene. If an individual is caught holding a smoking gun over a body, it indicts that suspect. However, if no smoking gun is found, the suspect might still be guilty (Punton and Welle 2015). Failing a smoking gun test somewhat strengthens rival hypotheses. For example, Vi staff credit ICRAF scientists as a source of new information on a topic; it is sufficient if corroborated by several sources, but not necessary, as it could be possible for ICRAF research to be transmitted to Vi without it being explicitly credited.

- A **"doubly decisive"** test consists of evidence that is both necessary and sufficient to confirm a hypothesis. Passing a doubly decisive test eliminates all other hypothesis (Collier 2011). For example, there is evidence that (a) ICRAF made substantive contribution to developing an agroforestry practice; (b) trained Vi staff on the application of the practice; and (c) Vi staff promoted the practice in its operational areas and cite ICRAF research as a source of the technical knowledge.

Table 3.2.1: Key Process Tracing Tests

Test	Description	Example
Straw in the Wind	Neither necessary nor sufficient	Vi staff state that it was likely ICRAF that developed a practice it promoted
Hoop Test	Necessary, but not sufficient	ICRAF undertook research related to a practice Vi promoted
Smoking Gun	Sufficient, but not necessary	Core Vi training manual credits ICRAF to the development an AF practice it covers
Doubly Decisive	Necessary and Sufficient	Evidence that ICRAF research contributed significantly to AF practice, with evidence of Vi promotion

Implementing the process tracing 'sub-study' involved interviewing 47 actors at ICRAF, Vi, the Regional Land Management Unit (RELMA), the Kenya Agricultural

and Livestock Research Organization (KALRO), and the Swedish University of Agricultural Sciences (SLU), as well as undertaking a comprehensive literature review of relevant research and extension materials over a period of approximately six months (LePage Morgan 2017).

3.3 Farmer Group Identification and Sampling

Following the village matching exercise described above, an advance team of enumerators was sent to each village ahead of the survey administration team to prepare lists of farmer group members. Our sampling frame required that all respondents in both treatment and comparison villages pass the following screening criteria:

- Must be a member of a group that was formed in 2008/09 or earlier
- Must have been an active member of that group since 2008/09 or earlier
- Household must have existed in 2007 or before
- Household must have been farming the same main parcel of land from 2007 to the present

Lists of all farmer group members active since at least 2008/9 were constructed in each of 121 sampled villages, and efforts were made to randomly sample 12 women and 12 men from these lists.

These screening criteria were used, in part, to mimic Vi's selection process when it engaged with the villages in the programme area during the early stages of its programme. Recall from Subsection 2.1 that it targeted pre-existing farmer groups, rather than mobilizing new ones. Hence, by identifying all farmer groups that had been active in both the programme and comparison villages at this time, coupled with sampling and interviewing their members that had also been so active, we sought to set up a fair comparison. In the programme area, interviewing members of all pre-existing famer groups, regardless of whether they actually engaged with Vi, ensured that we did not simply compare a unique set of farmer groups with a more general set in the comparison area. Similarly, identifying and interviewing all groups and members that had been active in the early stages of the programme period in the comparison area enable the identification of a set of farmers groups that would have been offered the opportunity to participate in Vi's programme had Vi gone to these villages and followed a recruitment approach similar to that used in the programme area.

In the event, the sampling team contacted farmer group leaders in the selected villages and requested lists of members who met the above screening criteria. Once the lists from each group in the village were assembled, the sampling team randomly selected 12 female respondents and 12 male respondents from each village, when sufficient numbers of females and males were available in the farmer groups' membership. However, male participation was at times insufficient, leading to greater proportion of females being sampled. We followed this approach, given our interest in testing for differential programme effects between female and male farmer group members. The sampled respondents were then informed and mobilized by the group leaders, in an attempt to increase the likelihood that they would be present and available for interviews upon the survey administration team's visit to the village.

3.4 Data Collection and Analysis

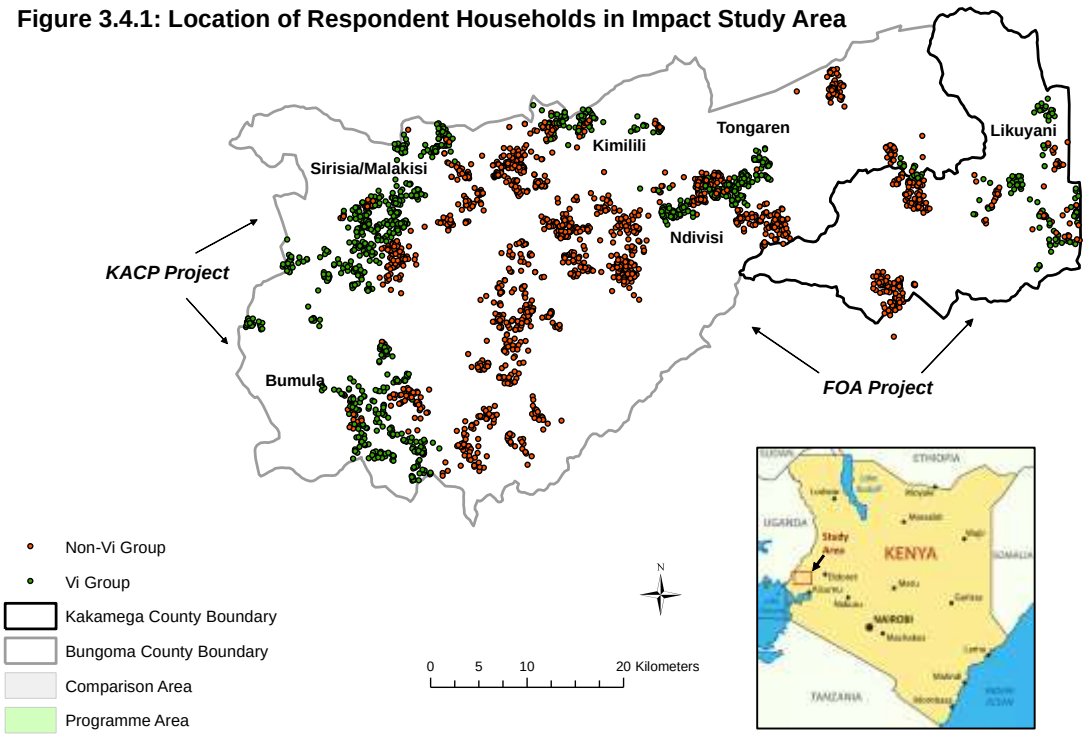
Based on (a) the above formative qualitative work; (b) the theory of change presented in Subsection 2.2; (c) a more in depth understanding of Vi's programme; and (d) discussions within our team, a draft survey instrument was developed using the Open

Data Kit (ODK) platform. This draft instrument was then reviewed with Vi program staff in Bungoma County, followed by piloting it with four smallholder farmers residing in this same county but outside of the impact study area. It was revised thereafter, and a team of 24 enumerators were trained to carry out its administration for a period of three days, with the assistance of a technical officer from the Kenya Agricultural and Livestock Research Organization (KALRO). This training included a practical exercise where each enumerator interviewed a farmer, followed by another extensive review of all the questions. Throughout the enumerator training programme, the survey instrument was iteratively refined. The enumerators were recruited from a pool of 306 applicants, from which 58 were selected for interviews and 24 finally shortlisted for training. Efforts were made to have an even number of female and male enumerators, given that 12 female and 12 male farmers were to be interviewed in each village.

Despite challenges experienced in the field administering the survey, a quality dataset comprising of 2,797 observations was derived.

The survey was carried out in the 121 programme and comparison villages from August 4 to October 1, 2016. This included an administration of a questionnaire to capture information on both the respondent, e.g. their age, educational status, and farmer group participation, and their household, e.g. educational status of other household members, baseline livelihood pursuits, and asset ownership. The last part of the interview process involved a visit to the household's main farming parcel (*shamba*). Here, the respondent was asked questions about this main parcel, e.g. tenure arrangements and size. The enumerator then visited each plot within the parcel, while making observations (e.g. tree presence), asking specific questions about the plot both at the present and during the baseline period (e.g. types of crops grown), and taking GPS coordinates.

Figure 3.4.1: Location of Respondent Households in Impact Study Area



Given challenges of (a) identifying existing farmer groups in each village that had been active since the baseline period; (b) ascertaining whether they are affiliated with Vi; and (c) subsequently profiling and sampling members who had been active since the

baseline period, this exercise was much more involving and time consuming than expected. It was particularly difficult for the sampling team to keep ahead of the survey team throughout much of the exercise. In addition, despite efforts to mobilize sampled farmer group members prior to the arrival of the survey team in the village in question, many were nevertheless difficult to locate, thereby making it challenging for the enumerators to reach their daily survey quota of four households each. That said, the dedicated efforts of the survey supervisors and enumerators paid off. In the end, 432 farmer groups operating in 121 villages (making up 1,450 and 1,410 households in the program and non-program villages, respectively) were interviewed. Of the households in the programme villages, 1,093 (75.38%) are members of farmer groups that participated in Vi's programme from the baseline period onwards. The total number of households that were successfully interviewed exceeded the target of 2,160 households by over 32%. Figure 3.4.1 presents the study area and the location of the respondents' homes.

During the data collection exercise, the data were downloaded periodically from a dedicated Internet site hosted by Ona and checked for survey administration errors. At the end of the survey exercise, the collected data were imported into Stata for cleaning, variable construction, and analysis. For continuous measures, outliers were addressed by trimming them to the 1 and 99 percentiles, particularly where data entry errors were not clearly identifiable. During this process, it became apparent that 63 interviewed farmer group members did not meet the screening criteria presented in Subsection 3.3, so were dropped from the dataset, reducing the total sample size to 2,797. A pre-analysis plan was also prepared following the International Initiative for Impact Evaluation's (3ie) Registry for International Development Impact Evaluations (RIDIE) format, uploaded onto its site, and subsequently formally accepted. See here for the publically accessible registration.

3.5 Baseline and Time Invariant Respondent and Household Covariate Balance

Several relevant baseline and time invariant differences exist between farmer group members and households in the programme and comparison areas.

The objective of the village geospatial and secondary data matching exercise was—ultimately—to help achieve a fair, unbiased comparison between the households of the programme and comparison areas. Consequently, it is important to assess how well the two groups of households are actually balanced across relevant baseline and time invariable variables (herein covariates). The full set of 46 covariates is presented in Annex 2, while Table 3.51 presents only those found to be statistically significant at a 90% level of confidence or greater (with eight out of the 46 correlated at the 95% level net of VSZ). As is clear, the farmer group members in the programme area are slightly more likely to be female and about 6% less likely to head their particular households. Moreover, while they are more likely be technically skilled, they are also 8% less likely to own their respective household's main farming parcel outright. Programme area households were also more likely to have had reared livestock and have one or more members in formal employment in 2007. They are further more likely to be elderly headed, have had soils richer in organic matter at baseline, and reside further from the main tarmac road. Finally, households of the programme area have slightly fewer children and lower levels of education, particularly when all adults in the household are taken into account.

While by no means extreme, these differences are both correlated with many of the study's outcome measures and were theoretically selected given their potential to

influence the status of these outcome measures over time. Hence, we included the all covariates presented in Table 3.5.1 significantly correlated with our programme area dummy (p<0.1 and net of VSZ) in all our models used to estimate the effects of Vi's programme. We did not control for formal land title and elevation, given that all our models controlled by our VSZ dummies and, hence, already netted out the difference.

Table 3.5.1: Covariate Comparison of HHs in Programme and Comparison Areas—Statistically Significant Differences Only

Characteristic	Program Mean	Non-program Mean	Difference (raw)	Difference (net of county)	Difference (net of VSZ)
Respondent Female	0.63	0.60	0.031*	0.082*	0.082*
			(1.70)	(1.69)	(1.69)
Respondent is Head	0.48	0.54	-0.057***	-0.14***	-0.14***
			(-3.01)	(-3.02)	(-3.03)
Respondent is Spouse of Head	0.41	0.37	0.038**	0.099**	0.100**
			(2.06)	(2.06)	(2.07)
Respondent has Tech. Skills	0.06	0.04	0.020**	0.18**	0.18**
			(2.30)	(2.31)	(2.29)
HH reared livestock in 2007	0.62	0.57	0.055***	0.14***	0.14***
			(2.97)	(2.97)	(2.92)
HH member employed in 2007	0.16	0.13	0.030**	0.13**	0.12**
			(2.24)	(2.21)	(2.08)
Head is 60 or older	0.33	0.29	0.043**	0.12**	0.12**
			(2.46)	(2.46)	(2.42)
HH had formal title to main parcel (07)	0.35	0.32	0.030*	0.082*	0.078
			(1.67)	(1.66)	(1.58)
Respondent owned main parcel (07)	0.46	0.54	-0.084***	-0.21***	-0.21***
			(-4.44)	(-4.44)	(-4.40)
Number of children in HH	3.56	3.71	-0.15**	-0.15**	-0.15*
			(-1.99)	(-1.99)	(-1.91)
Highest years of educ. of any adult in HH	10.72	10.94	-0.23*	-0.23*	-0.23*
			(-1.86)	(-1.87)	(-1.92)
Estimated 2007 soil organic car. (plot avg.)	24.38	22.97	1.42***	1.40***	1.36***
			(4.63)	(5.12)	(6.12)
Elevation (hh level)	1559.85	1529.87	29.6***	29.4***	-0.027
			(3.66)	(4.09)	(-0.61)
HH distance from tarmac road (km)	3.08	2.96	0.12*	0.12*	0.12*
			(1.67)	(1.68)	(1.77)
Observations	**1411**	**1386**	**2797**	**2797**	**2797**

z/t statistics in parenthesis; VSZ=Village Sampling Zone; * p<0.1, ** p<0.05, *** p<0.01
Probit regression used for net of county and VSZ differences, so coefficients are not directly interpretable, only the *t/z*-statistics

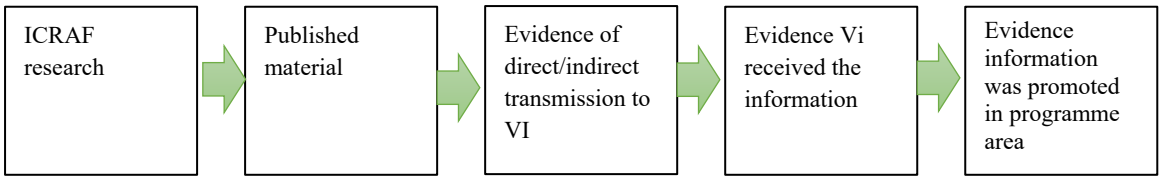
4. Empirical Results

In this section, we present the study's empirical results following the theory of change presented in Subsection 2.2. We start with a review of the key findings of the study's process tracing component that interrogated the causal linkages between ICRAF's research and the agroforestry practices and tree and shrub species Vi encouraged the participating farmer groups to take up. We then examine—through factual analysis—the veracity of the second and third key steps in this theory of change: that the members of these groups substantively participated in Vi's programme and appropriately took up (i.e. adopted) the promoted agroforestry practices and trees/shrubs. A presentation of the results of our counterfactual analysis then follows, which includes (a) estimations of the effects of Vi's programme on the theory of change's intermediate and final impacts; and (b) the extent to which the data are consistent with several hypothesized mechanisms pertaining to how the study's estimated downstream effects may have come about.

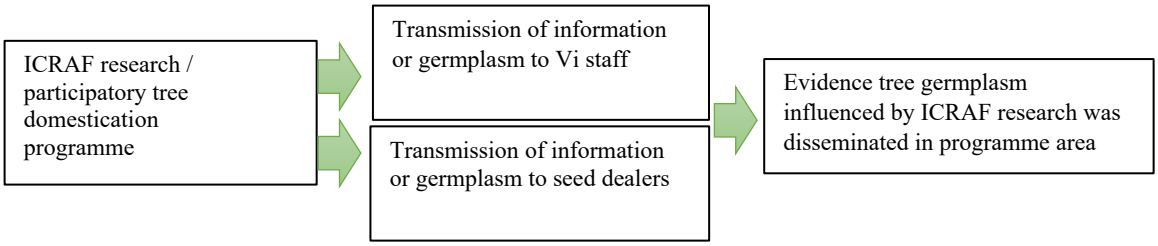
4.1 Connections between ICRAF/CGIAR research and Vi's programme

As explained in Subsection 3.2, a member of our research team carried out a comprehensive historical assessment of the linkages between ICRAF's research on the one hand and Vi's programme on the other. The particular evidence chains that were examined were as follows for specific agroforestry practices and tree/shrub germplasm, respectively.

Agroforestry Practice Evidence Chain:



Tree/shrub Germplasm Evidence Chain:



The results of the study are summarized in Table 4.1.1, while the full paper can be found here. As is clear, there is doubly decisive evidence linking ICRAF's research and Vi's promotion of (a) *Calliandra calothyrsus* and *Sesbania sesban* as protein-rich fodder; (b) both herbaceous and woody improved fallows; and (c) alley cropping for erosion control and soil fertility. There is also some evidence that the tree/shrub germplasm promoted by Vi in the programme area was informed by ICRAF's participatory tree domestication programme. However, this only passes the 'hoop' test. The evidence is less clear on whether Vi's promotion of understory multipurpose

trees or general soil and water conservation practices were significantly influenced by ICRAF's historical research efforts.

Table 4.1.1: Process Tracing Linking ICRAF Research to Vi Programme Summary Table

Key Highlights	Evidence of research contribu-tion?	Evidence of transmis-sion to Vi?	Evidence of promotion in field?	Hypothesis: ICRAF informed promoted practices
Fodder Shrubs				
• In the 1990s, ICRAF led fodder shrub trials with several breakthroughs, e.g. identifying promising fodder shrub species and establishing the substitution rate of *Calliandra calothyrsus* for dairy meal. • With a few exceptions, Vi's training programme on fodder shrubs includes recommendations specifically developed by ICRAF.	Doubly Decisive	Doubly Decisive	Doubly Decisive	Confirmed
Improved Fallows				
• ICRAF research on improved fallows showed that 1-4 year fallow rotations of leguminous trees can increase maize yields. • Vi's training programme in the programme area promotes species & techniques based in ICRAF's research.	Doubly Decisive	Doubly Decisive	Doubly Decisive	Confirmed
Alley Cropping				
• Alley cropping was developed as a technology by IITA in the 1970s. • ICRAF, IRRI, ILCA/ILRI, and IITA research on alley cropping in the 1980s & 1990s led to a better understanding of its limitations. • Vi's training programme in the project areas reflects the technical recommendations that came out of this research.	Doubly Decisive	Doubly Decisive	Doubly Decisive	Confirmed
Understory Multipurpose Trees				
• Vi promotes a practice of interspersing 'long-term trees' (such as *Grevillea robusta*) with 'short-term trees' (such as *Sesbania sesban).* • This practice is credited to ICRAF trainers, but this study was unable to locate ICRAF research on the topic.	Straw in the Wind	Hoop Test	Hoop Test	Unconfirmed
Soil and Water Conservation (SWC)				
• ICRAF's primary research has been on trees and shrubs, not specifically on SWC. • ICRAF made some research contributions in this area, including developing 'Natural Vegetative Strips' (NVS), and studying the roles of leguminous trees on contour bunds. • Vi's primary source of information on SWC was the organization RSCU/RELMA, which ICRAF hosted from 1994 -2006.	Hoop Test	Hoop Test	Hoop Test	Unconfirmed
Germplasm Supply				
• ICRAF's participatory tree domestication programme began in the mid-1990s, which focused on training local farmers, extension agents and seed dealers to identify desirable traits in trees, and to collect, germinate, and raise high-quality germplasm. • Vi staff received information on seed technology from ICRAF from the late 1980s onward. • Seed dealers who supply Vi with germplasm report receiving training on seed technology from both ICRAF and Vi. • Some quantities of *Calliandra germplasm* has been sold directly from ICRAF stands in the Maseno area to Vi. • Since 2007, ICRAF has developed technologies in nursery management by these have not been transmitted to Vi.	Hoop Test	Hoop Test	Hoop test	Confirmed

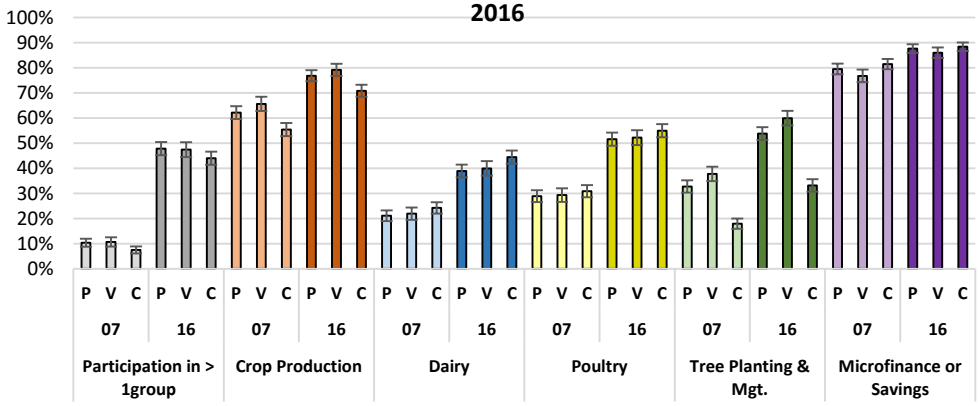
4.2 Vi Programme Exposure

Adequate exposure among the targeted farmer group members to Vi's programme is an essential precondition for the realisation of its expected outcomes and impacts.

A key assumption underpinning our theory of change for Vi's programme is that agroforestry was substantively promoted among the farmer groups that Vi targeted. To investigate the veracity of this assumption, the survey respondents were asked whether they had: (a) participated in one or more groups over the previous three years and back in 2007, including the number and their specific foci; and (b) been trained and/or received extension support in specific areas over the past three years and back in 2007, including the number of times and in which specific areas. Unfortunately, due to a programming error made in the ODK survey form's extension support module, we failed to capture the specific types of extension support the interviewed farmers received in both time periods. Consequently, the following presents greater detail on group participation and training and less on the receipt of extension support.

Figure 4.2.1 presents highlights on the participation of the respondents in groups, while greater detail and the results of statistical tests comparing the programme and comparison areas are presented in Annex 3. Recall from Subsection 3.3 that all interviewed farmer group members from both the programme and comparison sites had to belong to one or more farmer group from 2007 onwards. However, we created a variable indicating whether the respondent participated in more than one group. As indicated in Figure 4.2.1, approximately 10% of the respondents had participated in more than one group during the baseline period, with this percentage increasing to nearly 50% in 2016. While the differences between the programme area and Vi respondents on the one hand and those of the comparison sites on the other are statistically significant in favour of the former in both time periods, they only differ by a few percent points.

Figure 4.2.1: Reported Group Participation by Group Focus, 2007 & 2016



P = Programme Area; V= Vi Group; C = Comparison Area

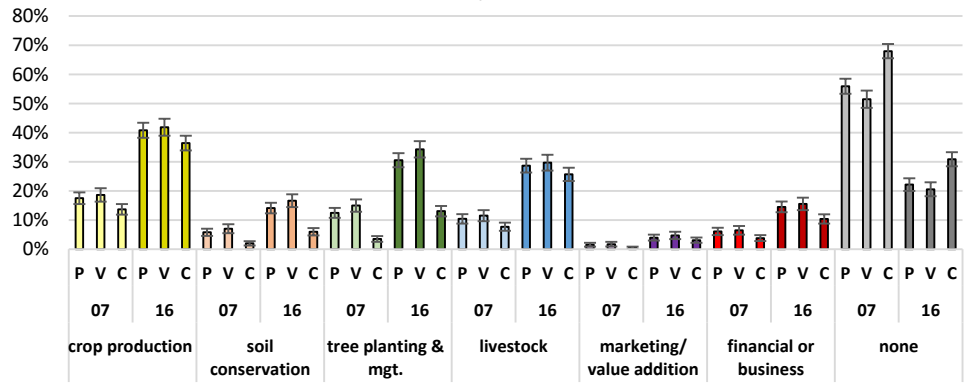
The foci of the groups during both time periods are particularly revealing. Participation in groups focusing on crop production and microfinance/savings is relatively high in both time periods for both the programme and comparison areas, albeit slightly higher for the former. Participation is relatively lower in the other areas. This result is particularly surprising for tree planting and management for the Vi group interviewees, in particular: While the difference vis-à-vis those in the comparison area

is large and highly statistically significant, 40% of VI group members did not report this distinguishing area of work as being a key focus of their respective groups.

Given that one of Vi's primary agroforestry promotional method is training, Figure 4.2.2 further presents more surprising results. While the receipt of significant training had increased in several areas from the baseline period, only 34% of Vi group members reported both being trained in tree planting and management and having had implemented this training to at least a medium extent in the past three years. While this figure increases to 50% when reference is just made to training, it is clear that large numbers of Vi group members did not receive any training in tree planting and management, either via Vi's programme or any other. Nevertheless, the receipt of such training was significantly greater than that received by farmer group members in the comparison area.

While much higher than households in the comparison area, only 60% of the Vi group interviewees reported tree planting and management as being a key focus of their groups and only 50% reported having had been trained on this topic in the last 3 years.

Figure 4.2.2: Receipt and Significant Implementation of Training in Last 3 Years, 2007 and 2016



P = Programme Area; V= Vi Group; C = Comparison Area

Figure 4.2.3: Breakdown of AF Training Recieved in Last 3 Years

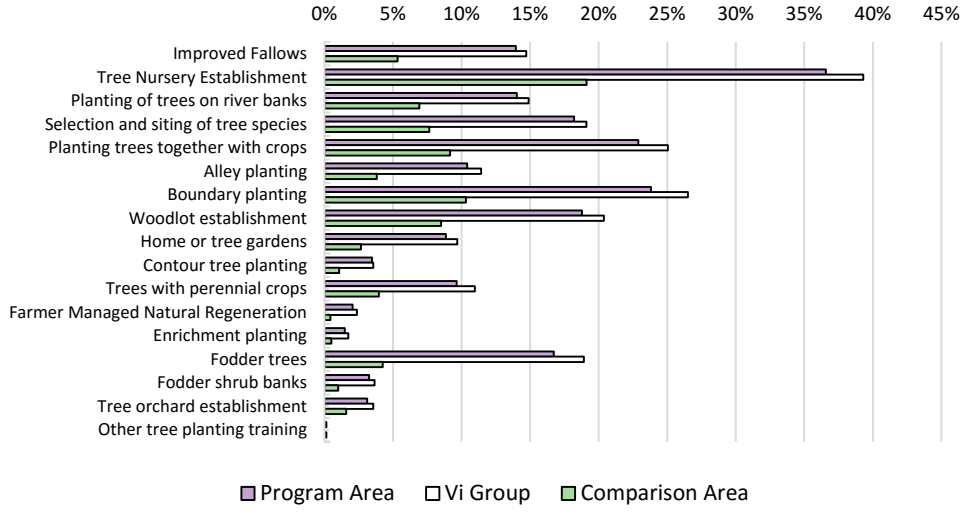
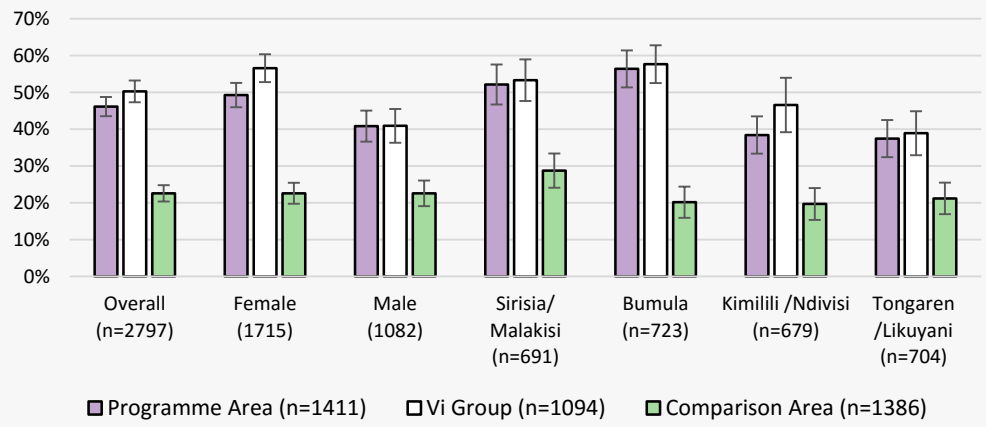


Figure 4.2.3 presents a breakdown of the specific types of tree planting and management training that was reported as having had been received during the previous three years. Most members in the programme areas reported that they had been trained in tree nursery establishment, followed by boundary planting and integrating trees within crop fields. Figure 4.2.4 disaggregates the receipt of tree planting and management by gender and VSZ. It is clear that more women were trained than men and that Vi group members in the Sirisia/Malakisi and Bumula VSZs received more training than in the other two VSZs.

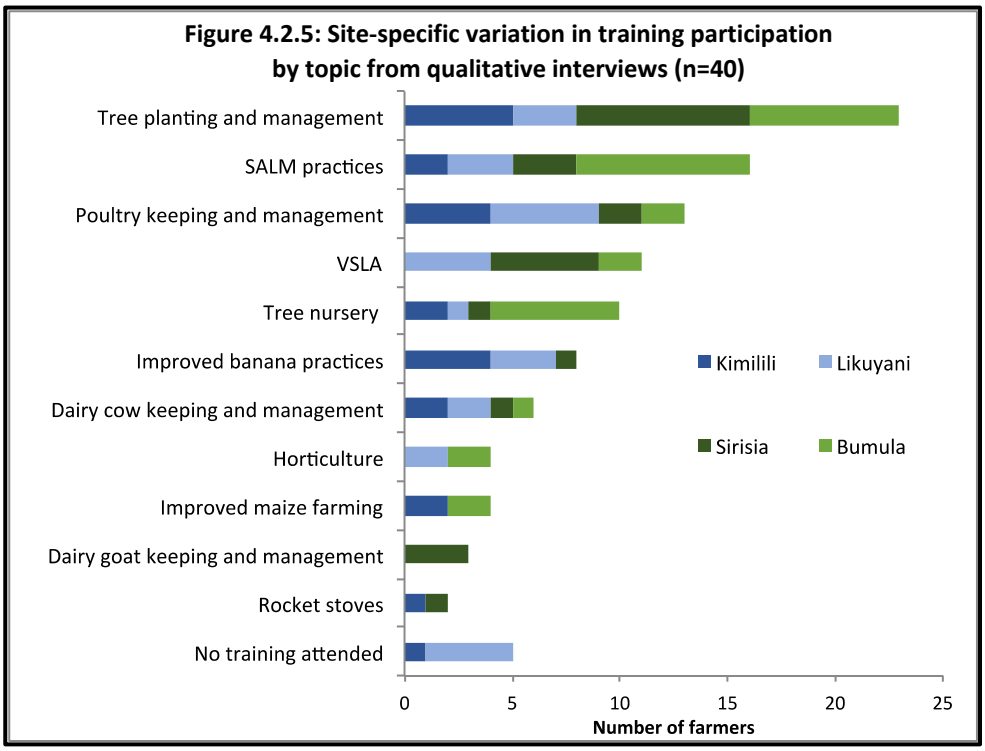
Figure 4.2.4: Reported Receipt of Training in Tree Planting & Mgt. in Past 3 Years



Both the study's quantitative and qualitative components evidenced gender and site specific variation in the receipt of agroforestry related training.

Our qualitative data complements the above story: While the most commonly cited, only 23 out of the 40 farmers (14 out of the 23 women and nine of the 17 men) reported have had received training in tree planting and management. Similar to the quantitative findings, the highest numbers are in the Bumula and Sirisia sites. See Figure 4.2.5.

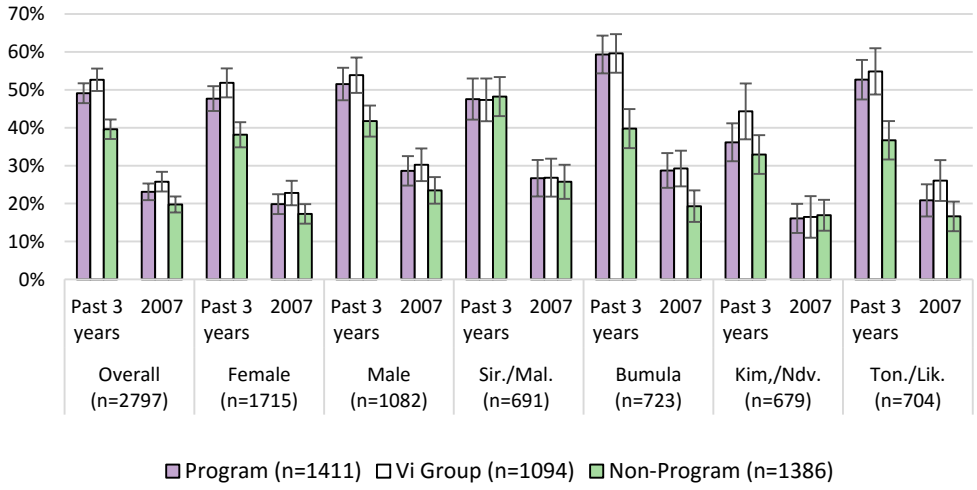
Figure 4.2.5: Site-specific variation in training participation by topic from qualitative interviews (n=40)



The theory of change's precondition of adequate programme exposure was only partially met.

While we do not have data on the specific types of extension support received, Figure 4.2.6 reveals that only about 50% of Vi group members reported having had received extension support in any area for the last three years. This ranges from 60% in Bumula VSZ to 44% in Kimilili/Ndivisi VSZ, with the differences between female and male respondents being about the same. Moreover, only 14 out of the 40 farmers interviewed during the qualitative follow-up exercise stated that they had experienced a home monitoring visit by Vi, with the majority (nine) residing in the Bumula VSZ.

Figure 4.2.6: Reported Receipt of Extension Support in Last 3 Years and in 2007



From the both the study's quantitative and qualitative findings, there is evidence that exposure to agroforestry promotion was significantly greater in the programme area in general and among Vi group members in particular. However, it is also clear that such exposure was not substantive for many Vi group members. We will now examine the extent to which this trend is similar for the uptake of the particular agroforestry practices and tree and shrub germplasm promoted by Vi's programme.

4.3 Adoption of Promoted Agroforestry and Related Practices and Tree Germplasm

In this section we examine the extent to which the agroforestry and other Sustainable Agricultural Land Management (SALM) components of Vi's programme were taken up (adopted) by the farmer group members in both the programme and comparison areas. Given the complexity and richness of both components, we present two indices to enable the data associated with each to be practically aggregated and analysed—an Agroforestry Adoption Index and another index for the other SALM practices that were promoted.

4.3.1 Agroforestry Practices and Tree Species

While both we and Vi staff recognize that there is 'no one size that fits all' for agroforestry, we worked with the latter to devise specific indicators to reveal the extent to which the promoted agroforestry practices and tree and shrub species had actually been taken up by the targeted farmers. Following several rounds of iteration, we came up with the 10 binary indicators presented in Figure 4.3.1. These are grouped under three dimensions: Practice Uptake; Intensity of Practice; and Tree Species. If a

household had significantly taken up Vi's promoted agroforestry practices and tree germplasm, we would expect to see trees and shrubs not only integrated into its farming plot(s) (i.e. where food and horticultural crops are grown) but also the harvesting of several tree products, i.e. fuelwood, timber, fruits and/or fodder. We would also expect to see tree-based natural resource management (NRM) techniques being applied, such as planting trees along contour lines interspersed with shrub species. Several other complementary agroforestry practices would also be expected to be present on the farm, such as fruit orchards, woodlots, and/or fodder banks.

To measure the uptake of Vi's programme, we construed a multi-dimensional 'agroforestry adoption' index specific for Vi's programme.

Figure 4.3.1: Agroforestry Adoption Index for Vi's Programme
(Dimensions and Indicators)

We would further expect that the intensity of these practices to be significant. We would expect to see, in particular, the presence of a relatively high density of trees on the adopting household's food and horticultural plots, together with other complementary agroforestry practices, coupled with significant income earned through the sale of the resulting products. Finally, we would expect to find evidence that the household had planted some of the 'signature' tree species promoted by Vi, ranging from leguminous shrubs through to more long-term exotic and native species.

Following Alkire and Foster (2011) but without discounting, we took the above 10 binary indicators and weighted them equally under each dimension to create an agroforestry adoption index specifically relevant for Vi's programme. It ranges from 0 to 1, with a score of 1 revealing that a household has surpassed the binary cut-off on all the 10 indictors and 0 if otherwise. Table 4.3.1 present mean values for both 2007 and 2016 versions of the index, as well as the difference between the two time periods and for each of the three dimensions.[3] This is for the sampled households in the programme and non-programme areas, as well as for two sub-groups of households within the former, i.e. Vi group members and other farmer group member households residing in the same villages Vi targeted. Table A4.1 in Annex 4 complements this by comparing each specific indicator separately.

[3] Each of the three dimensions was reweighted to fall on the same scale ranging from 0 to 1. This enables each to be compared with one another and the overall index.

Table 4.3.1: Comparison of HHs—Agroforestry Adoption Index and Dimensions

Binary Indicators		PA Mean	Vi Group Mean	Non-Vi in PA Mean	Non-PA Mean	PA vs. non-PA (dif.)	Vi vs. non-PA (dif.)	Vi vs. non-Vi in PA (dif.)
Overall Index	2007	0.15	0.15	0.13	0.11	0.035***	0.039***	0.017
						(5.67)	(5.91)	(1.55)
	2016	0.27	0.28	0.20	0.17	0.090***	0.11***	0.082***
						(12.17)	(13.77)	(5.98)
	Change	0.12	0.13	0.07	0.06	0.055***	0.070***	0.064***
						(8.80)	(10.38)	(5.44)
Dimension 1: Practice Uptake	2007	0.20	0.20	0.18	0.15	0.044***	0.049***	0.023
						(5.29)	(5.51)	(1.57)
	2016	0.31	0.33	0.25	0.23	0.082***	0.100***	0.077***
						(8.72)	(9.89)	(4.57)
	Change	0.11	0.13	0.07	0.08	0.039***	0.051***	0.054***
						(4.43)	(5.43)	(3.34)
Dimension 2: Intensity of Practice	2007	0.13	0.13	0.11	0.10	0.022***	0.026***	0.015
						(3.12)	(3.38)	(1.19)
	2016	0.22	0.24	0.17	0.15	0.069***	0.085***	0.071***
						(8.19)	(9.49)	(4.65)
	Change	0.10	0.11	0.05	0.05	0.047***	0.059***	0.056***
						(5.59)	(6.60)	(3.76)
Dimension 3: Tree Species	2007	0.12	0.12	0.11	0.08	0.039***	0.042***	0.014
						(4.86)	(4.94)	(0.93)
	2016	0.26	0.28	0.19	0.14	0.12***	0.14***	0.096***
						(11.49)	(12.69)	(5.04)
	Change	0.14	0.16	0.08	0.06	0.081***	0.099***	0.083***
						(8.93)	(10.20)	(4.92)
	Observations	**1411**	**1094**	**317**	**1386**	**2797**	**2480**	**1411**

t statistics in parenthesis; PA = Programme Area
* p<0.1, ** p<0.05, *** p<0.01

There are a number of noteworthy observations. The first is that, while the index scores are low in the 2007 period, they are slightly higher in the programme area in general and among the Vi group members in particular. This is the case across all three dimensions, indicating that the practice of agroforestry was likely already slightly higher in the programme area prior to the coming of Vi. However, given that recalled data were used to construct the 2007 index, we cannot rule out the possibility that recall bias was systematically different for the Vi group members, e.g. given their assumed greater knowledge of agroforestry. Nevertheless, the second observation is that the relative change from 2007 to 2016 is significantly greater for the programme area and Vi groups overall and across all three dimensions. There is evidence, therefore, that many households in the programme area took up what Vi promoted. The final observation, however, is that the average index scores for 2016 for these two groups is not particularly high at 0.27 and 0.28, respectively, revealing that this take up was not as intense as perhaps desired.

Figure 4.3.2 is another way of presenting the same data presented in Table 4.3.1. It is clear, again, that the gain in the index scores from 2007 to 2016 was considerably greater in the programme area in general and among Vi group members in particular. The gain was, in fact, greatest with respect to the index's tree species dimension. Indeed, Table A4.1 in Annex 4 reveals that this is largely driven by their relatively greater uptake of leguminous shrub species. In particular, the percentage of farm plots found with such species increased from 16% to 48% among Vi households, against 8% to 20% among those of the comparison area.

The relative gains made on the agroforestry index were considerably greater in the programme area in general and among Vi group households in particular.

Figure 4.4.3 decomposes the gains made with respect to the index by gender of farmer group member. While greater gains in the index were made for both programme area and Vi group households with male members, the *relative degree of change* over the their counterparts in the non-programme area was as equally great for households with female farmer group members.

Figure 4.3.2: HH Comparsion on Agroforesty Index: Programme Area, Vi Group Only & Non-programme Area by Weighted Dimension

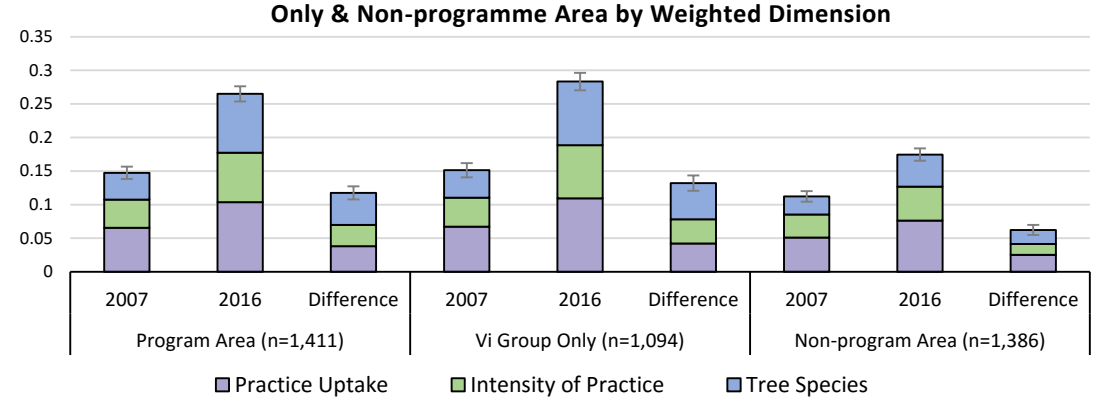


Figure 4.3.3: Relative Gains in Agroforesty Index: Programme Area, Vi Group Only & Non-programme Area by Gender & Weighted Dimension

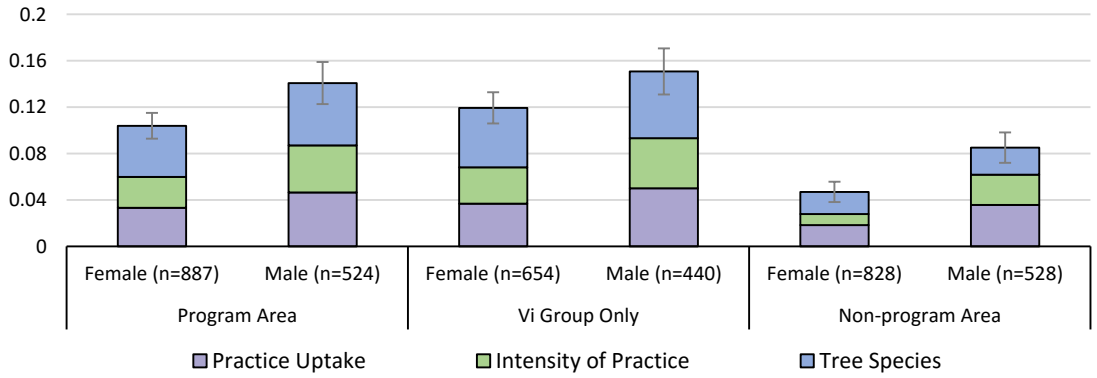


Figure 4.3.4: Relative Gains in Agroforesty Index by Weighted Dimension & Village Sampling Zone

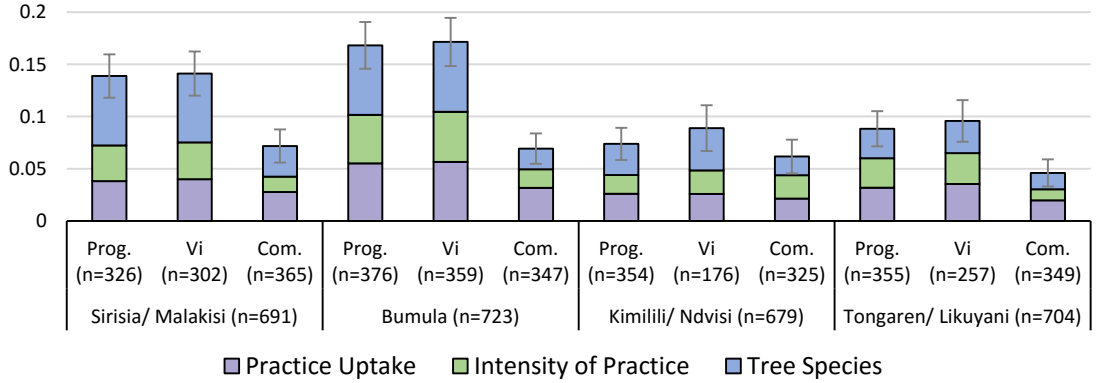


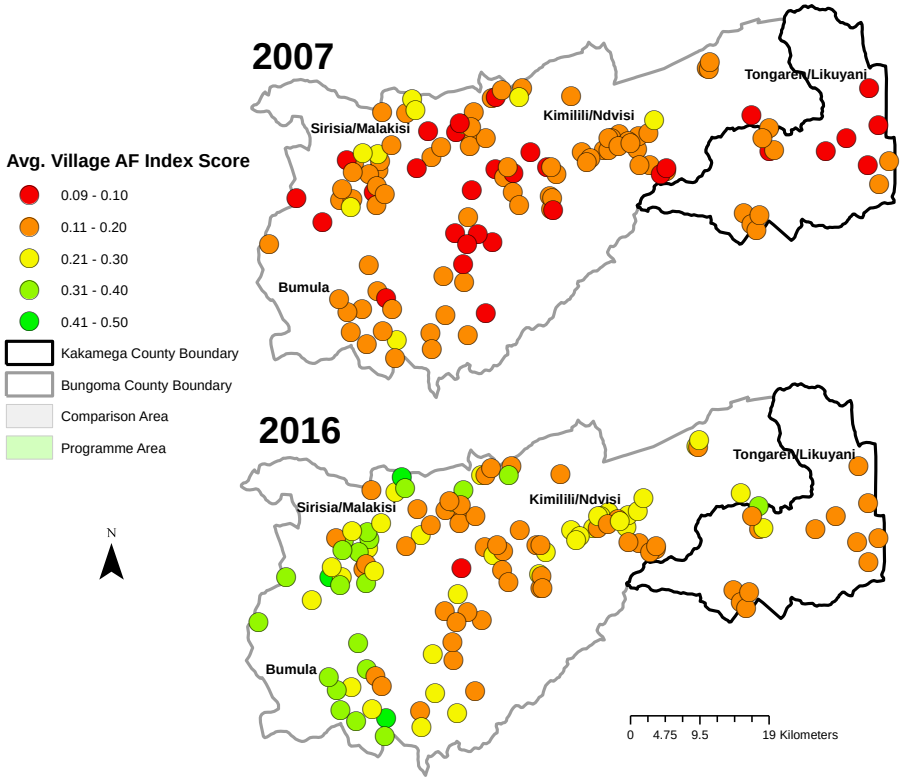
Table A5.1 in Annex 5 presents the results of statistical tests that examined the extent to which the overall average results for the index and its specific dimensions differ by geographic area and specific sub-groups (hereafter referred to as subgroup analysis). Here, it is clear that the results do not differ significantly for households with female and male members, nor by sex of household head. There is considerable variation among the four Village Sampling Zones (VSZs), however, with many of these differences being statistically significant. This variation is highlighted in Figure 4.3.4. The greatest and least gains took place in the Bumula and Kimilili/Ndivisi VSZs,

respectively. Indeed, the difference between households in the programme and comparison areas for the latter is statistically insignificant.

Figure 4.3.5 reveals this variation in the form of a map showing the spatial range of village average scores. While it is clear that agroforestry uptake took place in both the programme and non-programme areas, significantly greater uptake took place in the former, particularly in the Sirisia/Malakisi and Bumula VSZs.

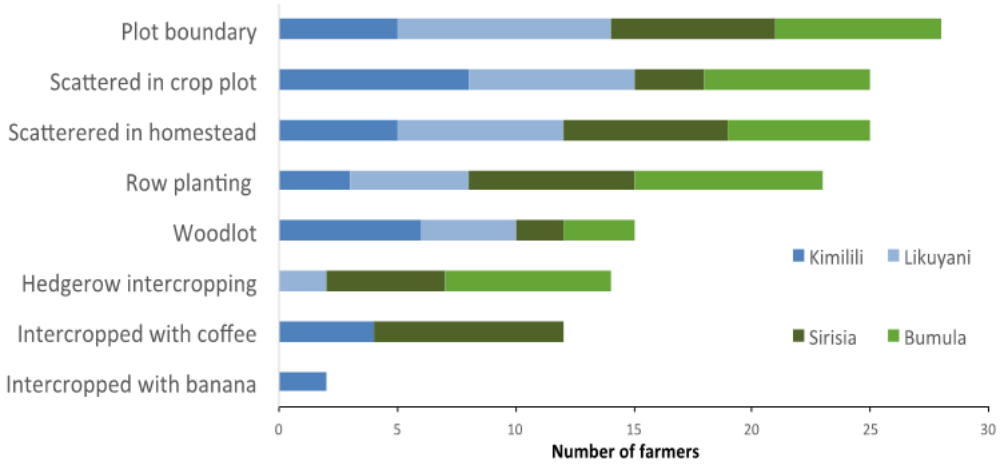
There was significant variation in the uptake of Vi's programme among the four VSZs and among dairy versus non-dairy producers.

Figure 4.3.5: Village Level Averages in Agroforestry Adoption Index 2007 and 2016



Our qualitative findings corroborates this spatial variation in the uptake of agroforestry practices. In particular, we documented eight types of agroforestry management practices across the study area (Figure 4.3.6). The most dominant practices observed on the 40 purposively sampled farms were tree planting along plot boundaries and interspersed trees within these plots and/or at the homestead. Approximately, one-third of farmers (14) planted at least one shrub species as hedgerows within their fields (alley cropping), with most coming from the Bumula and Sirisia sites. No hedgerows were identified in Kimilili, and they were only observed on two farms in Likuyani. Higher numbers of farmers (23) had planted rows of timber species along either contours or trash lines as promoted by Vi, but—yet again—this was mostly observed in the Bumula and Sirisia sites. Some practices, such as coffee intercropping were only found in the highland farming systems of Sirisia and Kimilili, the latter being the only place where farmers were also observed intercropping trees in banana plots. Finally, 15 farms had woodlots, mostly in the Kimilili site, and no fruit orchards were observed on any of the 40 purposively sampled farms, despite this being explicitly promoted by Vi.

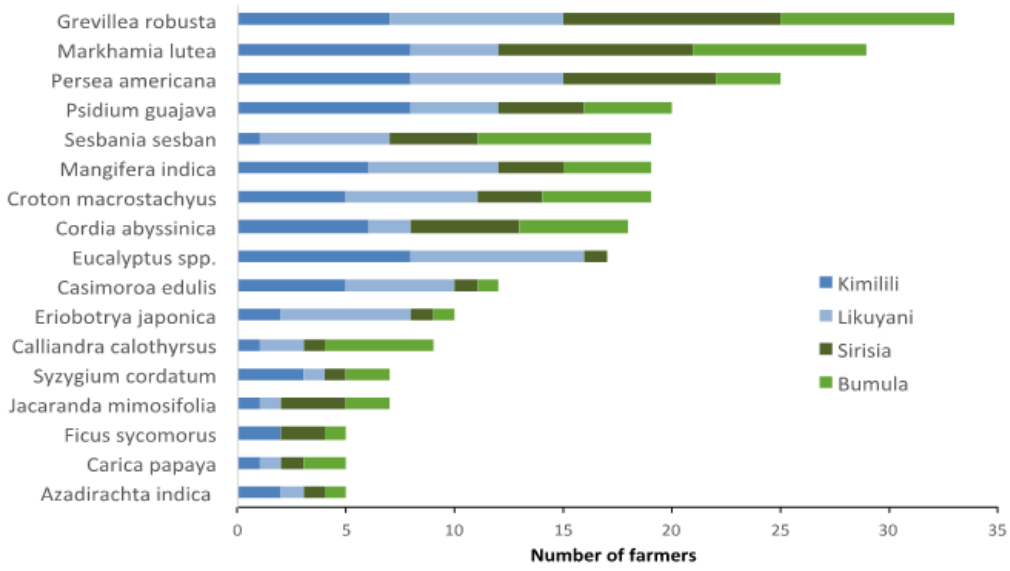
Figure 4.3.6: Geographic variation in agroforestry practices



Our qualitative work also revealed significant geographic variation in the uptake of agroforestry practices and tree/shrub germplasm.

Geographic variation was similarly observed in the types of tree and shrub species observed in the plots and homesteads of the 40 purposively sampled farms. The most dominant species is, unsurprisingly, *Grevillea robusta,* an exotic species native to coastal Australia and valued for its timber, closely followed by *Markhamia lutea*, a species native to East Africa with bright yellow flowers and valued for being relatively fast growing and terminate resistant (Figure 4.2.7). In addition to fruit trees (avocado and guava), the next most common species was *Sesbania sesban*—a high protein fodder shrub with a high percentage of foliage nitrogen—found on 19 farms, but with variation across the sites—Bumula (80%), Likuyani (60%), and Sirisia (40%) and Kimilili (one farm only). *Calliandra calothyrsus*, another high protein fodder shrub, was further found on about a quarter of the farms visited, though largely in Bumula. Amongst other trees promoted by Vi agroforestry, *Cordia africana* and *Croton macrostachyus* were grown by less than half of the farmers. Finally, Eucalyptus was found in Kimilili and Likuyani but not in Bumula and only on one farm in Sirisia. This is revealing, given that Vi actively discourages farmers from growing this 'thirsty' exotic species within or near crop land.

Figure 4.3.7: Geographic variation in tree species identified in farms & homesteads



The majority of the above identified species originated from seeds propagated by the interviewed farmers themselves, either through their own seed collections or using cuttings. In terms of external supply, Vi was clearly the largest source. However, it is noteworthy that, although 10 different species were sourced from Vi agroforestry, three species dominated this list: *Sesbania sesban* (16 farms), followed by *Grevillea robusta* (13 farms) and *Calliandra calothyrsus* (nine farms). Vi had provided *Grevillea* seedlings to just under a third of the farms visited, while One-acre Fund, another NGO operating throughout the impact study area, had supplied six farms, with two benefiting from both NGOs. Only three farmers had received fruit tree seedlings from Vi, while five farmers reported they are now propagating *Sesbania* through natural regeneration and transplantation.

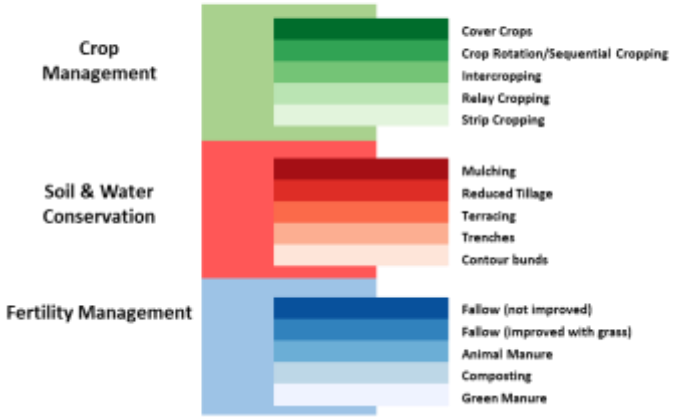
A particularly surprising sub-group difference related to the uptake of Vi's programme evident in the quantitative data is displayed in Table A5.1 of Annex 5. In particular, we hypothesized in our pre-analysis plan that dairy producers would experience greater impacts in terms of household income and asset accumulation, given the effects of tree fodder on bolstering and/or lowering the costs of milk production. However, for the uptake of Vi's overall programme at least, we found the opposite. As indicated in Table A5.1, while the index gains among dairy producers in the programme area are significantly greater than their counterparts in the comparison area, this same difference is nearly 60% greater among non-dairy producers.

4.3.2 Other Sustainable Land Management Practices

Vi's programme also promoted other sustainable land management practices, so it is also important to assess the extent to which these were taken up in both the programme and comparison areas.

As discussed in Subsection 2.1, Vi's programme also involved promoting other sustainable agricultural land management (SALM) practices that fall outside of a strict definition of agroforestry. Consequently, if the uptake of these practices was significant and significant differences are also found between the programme area and Vi groups on the one hand and the comparison area on the other on our study's socioeconomic and land health measures, it may prove difficult to disentangle the relative contribution each may have played.

Figure 4.3.8: Index for Other Sustainable Land Management Practices Promoted by Vi Agroforestry
(Dimensions and Indicators)



Plot specific data were therefore also captured on a total of 15 non-agroforestry SALM practices covered in Vi's SALM manual. Again, to compare the programme and

For the other non-agroforestry practices that Vi promoted, we created another complementary index.

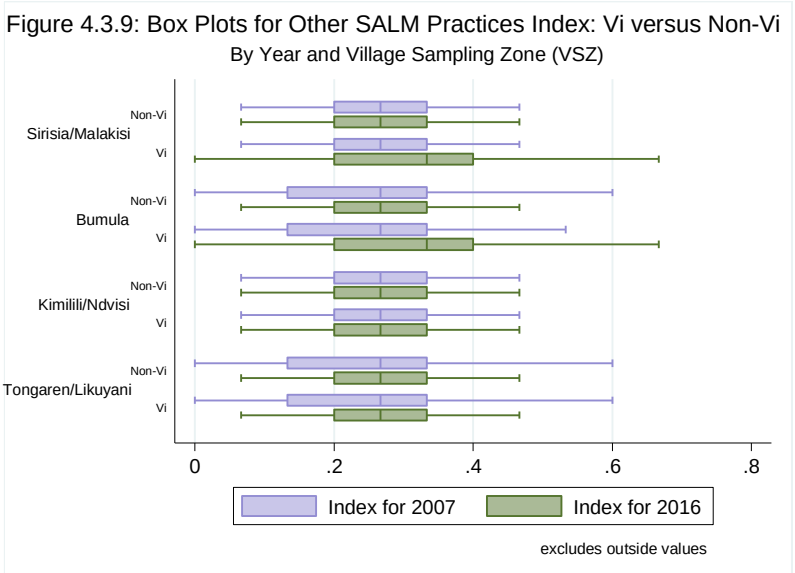
comparison areas against both in an integrated way, we grouped them into the three categories or dimensions presented in Figure 4.3.6 and created an 'other SALM practice index'. Like the AF Index, each practice is weighted equally under each dimension. The results are presented in Table 4.3.2, while a comparison of the groups against each of the 15 practice indicators is presented in Table A5.2 in Annex 5. For the overall index, the groups were at about the same level in the baseline period. However, we found modest, albeit statistically significant differences, for the 2016 index. The differences for the 2016 and 2007 indices are statistically insignificant when the programme and comparison areas are compared. This is also the case for the crop management and soil and water conservation dimensions, but the difference in favour of the programme area for the fertility management dimension is statistically significant. Moreover, the overall difference is significant when Vi households are compared with both (a) all households in the non-programme (comparison) area; and (b) non-Vi households in the programme area.

Table 4.3.2: Comparison of HHs—SALM Index and Dimensions

Binary Indicators		PA Mean	Vi Group Mean	Non-Vi in PA Mean	Non-PA Mean	PA vs. non-PA (dif.)	Vi vs. non-PA (dif.)	Vi vs. non-Vi in PA (dif.)
Overall Index	2007	0.26	0.26	0.25	0.25	0.0047	0.0049	0.0010
						(1.07)	(1.05)	(0.14)
	2016	0.29	0.30	0.27	0.28	0.010**	0.016***	0.024***
						(2.37)	(3.31)	(3.14)
	Change	0.03	0.04	0.02	0.03	0.0058	0.011***	0.023***
						(1.62)	(2.81)	(3.74)
Dimension 1: Crop Management	2007	0.50	0.50	0.52	0.52	-0.019**	-0.023**	-0.015
						(-2.17)	(-2.39)	(-1.01)
	2016	0.54	0.54	0.53	0.55	-0.016*	-0.013	0.010
						(-1.82)	(-1.45)	(0.72)
	Change	0.04	0.04	0.02	0.03	0.0037	0.0095	0.026**
						(0.51)	(1.20)	(2.13)
Dimension 2: Soil & Water Conservation	2007	0.09	0.09	0.07	0.07	0.014***	0.018***	0.016**
						(3.17)	(3.69)	(1.97)
	2016	0.11	0.12	0.08	0.10	0.014***	0.022***	0.034***
						(2.69)	(3.82)	(3.63)
	Change	0.03	0.03	0.01	0.03	-0.00031	0.0037	0.018**
						(-0.07)	(0.76)	(2.26)
Dimension 3: Fertility Management	2007	0.18	0.18	0.17	0.16	0.019***	0.019***	0.0020
						(3.10)	(2.98)	(0.19)
	2016	0.22	0.23	0.20	0.19	0.033***	0.039***	0.027**
						(5.07)	(5.60)	(2.45)
	Change	0.04	0.05	0.02	0.03	0.014**	0.020***	0.025***
						(2.52)	(3.33)	(2.62)
	Observations	**1411**	**1094**	**317**	**1386**	**2797**	**2480**	**1411**

t statistics in parenthesis; PA = Programme Area; * p<0.1, ** p<0.05, *** p<0.01

Figure 4.3.9: Box Plots for Other SALM Practices Index: Vi versus Non-Vi

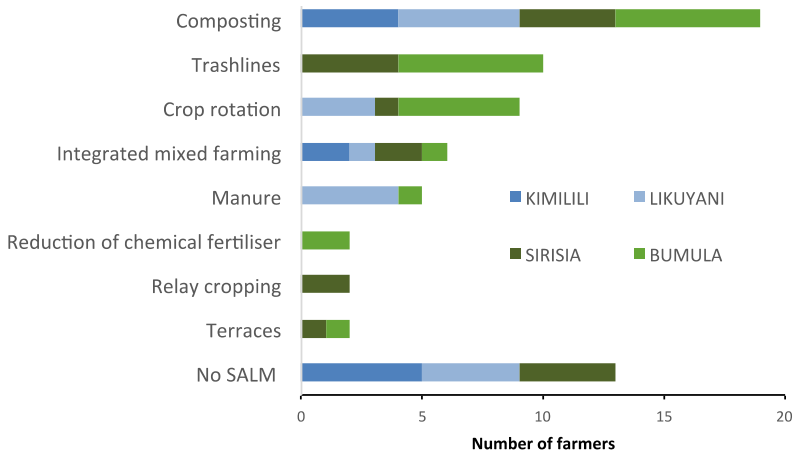


There are differences in the uptake of the other SALM practices among the VSZs as well. This is evident by examining the box plots presented in Figure 4.3.9. Here, it is clear that a number of Vi households that took up these practices over the life of the programme in the Sirisia/Malakasi and Bumula VSZs, which was not the case in Kimilili/Ndivisi and Tongaren/Likuyani.

The results of our qualitative work, again, corroborates these quantitative findings (Figure 4.3.10). Indeed, the uptake of the promoted other SALM practices was found to be low across the study area in general. However, the famers in the villages under KACP in the Bumula and Sirisia sites had a higher number of farmers adopting one or more of such practices. Composting for improvement of soil fertility was the most dominant practice, documented in all four locations by a little over half of the 40 interviewed farmers. Trash lines were the second most popular practice, but limited to the Bumula and Sirisia sites. The latter sites were also the only one where several farmers were found practicing relay cropping and terracing. All farmers interviewed in Bumula had at least one SALM practice in place. Improved soil fertility with better manure management was the dominant other SALM practice identified in Likuyani.

Both the quantitative and qualitative data reveal the uptake of other SLAM practices to be low overall, but with greater uptake among the households targeted by the KACP project.

Figure 4.3.10: Geographic Variation in the Uptake of Other SALM Practices across the Programme Area Found Through Qualitative Work

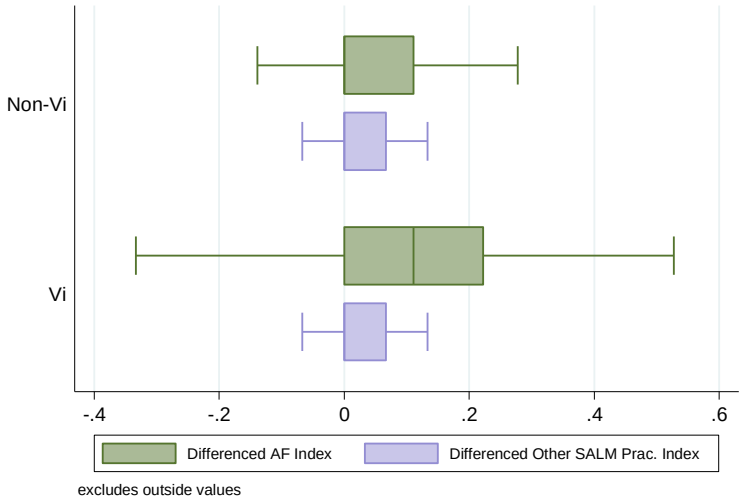


It is important to point out that, even though the uptake of the agroforestry practices promoted by Vi was not significantly intense overall, this uptake was clearly greater than that for the promoted other SALM practices. This is visually evident from the box plots presented in Figure 4.3.11. The changes for agroforestry uptake are both greater for all households overall and with respect to Vi affiliated households in particular

In summary, in this subsection we have seen that, while agroforestry in the programme area in general and among Vi affiliated households in particular was not particularly dramatic at least overall, it was significantly greater in comparison with the comparison area. It was also considerably more substantial than the other SALM practices that Vi promoted. In the Sirisia/Malakasi and Bumula VSZs, in particular, the gains in the Agroforestry Index were over twice that of their respective comparators. Moreover, if significant differences are found between programme and non-programme areas on the study's various outcome and impact measures, it is unlikely that this would have been driven by the uptake of the other non-agroforestry related SALM practices that Vi promoted, given that the differential uptake of these practices between the programme and comparison areas was minimal.

While Vi promoted various other SALM practices alongside agroforestry, this is unlikely to be a significant contributor to any of the identified downstream impacts.

Figure 4.3.11: Box Plots Comparing Changes in AF and Other SALM Practices Indices



In the next section, we will complement the preceding analysis of our Agroforestry Index by comparing the plots of the sampled households of both the programme and comparison areas against remote sensing derived estimates of tree cover. Indeed, if the integration of trees within these plots was significant among the Vi households, we should see this reflected in our analysis of the remote sensing data as well.

4.4 Remote Sensing Derived Tree and Vegetation Cover Estimates

Baseline and endline estimates of tree and vegetative cover, as well as measures of soil health, were derived by applying ground-truthed predictive models to 30x30m satellite imagery.

Through the GPS coordinates collected from the surveyed plots of each household's main parcel, measures of tree cover, vegetation cover, soil organic carbon (SOC), and soil erosion prevalence were estimated via remote sensing.[4] These measures are based on predictive models for mapping land health based on systematically collected field and lab measurements from approximately 150 Land Degradation Surveillance Framework (LDSF) sites across the global tropics. The LDSF consists of a set of methods for capturing data on various ecosystem health metrics, including soil condition, vegetation structure and cover, landform, land use, and land degradation (e.g. soil erosion). It is specifically developed for integrated landscape analysis, using a spatially balanced hierarchical sampling design based on 10 by 10 kilometre (100 km^2) sites (or landscapes), with each site consisting of 16 clusters, each 1 km^2, with 10 randomly generated 1,000 m^2 plots per cluster (Vågen et al. 2013) .

Table 4.4.1 presents the overall results for the plot level tree and vegetative cover estimates, respectively. While the plots surveyed in the programme area are estimated to have had approximately 2.2% greater tree cover, on average, in the baseline period (p=0.000), they also gained approximately an additional 3% over those in the comparison area during the programme period, with estimates for Vi group households being nearly 4% (2SLS coefficients). As is clear, the results are robust to various model specifications.

[4] The point estimates take into account the size of the fields where the GPS coordinates were captured. To mitigate the possibility of satellite imagery pertaining to areas outside of the sampled fields convoluting the results, estimates closer to where the GPS points were taken were given greater weight.

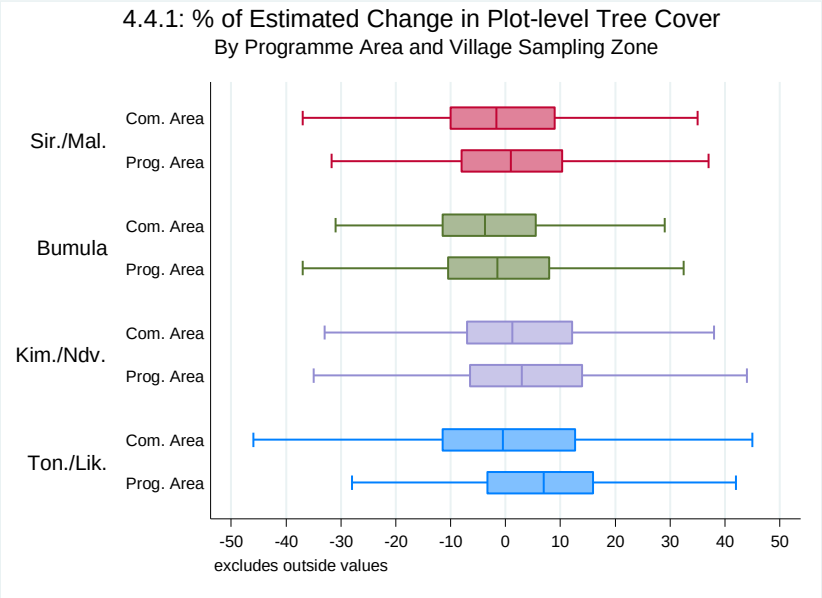
TABLE 4.4.1: Remote Sensing Derived Tree Cover & Green Vegetative Cover Estimates

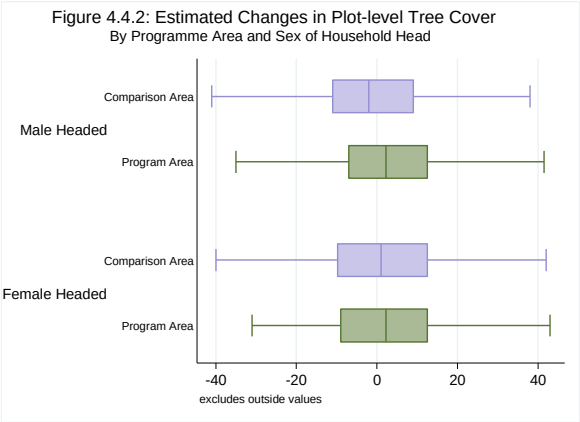
	Est. % of plots of main parcel with sig. tree cover				***Est. % of veg. cover across plots of main parcel***			
	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.
Raw results								
Sample Mean	56.68	1.23	56.66	1.15	59.86	-3.82	59.73	-3.98
Program Mean	59.52	2.99	59.53	2.93	59.99	-3.35	59.82	-3.55
Non-Program Mean	53.79	-0.56	53.73	-0.68	59.73	-4.30	59.64	-4.43
Unadjusted dif.	5.73***	3.54***	5.79***	3.61***	0.26	0.95**	0.18	0.88**
	(0.40)	(0.57)	(0.42)	(0.61)	(0.38)	(0.38)	(0.38)	(0.39)
Observations	**2788**	**2785**	**2757**	**2754**	**2790**	**2790**	**2759**	**2759**
Raw results								
Sample Median	56.50	0.50	56.50	0.50	59.99	-4.87	59.78	-4.91
Program Median	59.00	2.25	59.00	2.00	60.45	-4.42	60.45	-4.55
Non-Program Median	53.83	-1.50	54.00	-1.00	59.33	-5.25	59.23	-5.42
Unadjusted dif.	5.33***	3.75***	5***	3***	1.10**	0.83*	1.19**	0.88**
	(0.54)	(0.69)	(0.52)	(0.76)	(0.49)	(0.44)	(0.49)	(0.44)
Observations	**2788**	**2785**	**2757**	**2754**	**2790**	**2790**	**2759**	**2759**
OLS (ITT)								
Coefficient	4.99***	2.97***	5.11***	3.10***	-0.012	0.66	-0.097	0.56
	(0.61)	(0.78)	(0.63)	(0.81)	(0.56)	(0.58)	(0.56)	(0.59)
Observations	**2784**	**2784**	**2753**	**2753**	**2790**	**2790**	**2759**	**2759**
IPWRA (ITT)								
Coefficient	4.77***	2.91***	4.90***	3.04***	-0.037	0.64**	-0.097	0.53*
	(0.37)	(0.56)	(0.39)	(0.60)	(0.29)	(0.29)	(0.29)	(0.30)
Observations	**2784**	**2784**	**2753**	**2753**	**2790**	**2790**	**2759**	**2759**
Robust Reg. (ITT)								
Coefficient	4.91***	2.70***	5.00***	2.71***	0.091	0.68**	0.026	0.60**
	(0.38)	(0.58)	(0.40)	(0.62)	(0.26)	(0.27)	(0.27)	(0.27)
Observations	**2784**	**2784**	**2753**	**2753**	**2790**	**2790**	**2759**	**2759**
Nearest Neighbour 1 to 1 Matching (ITT)								
Coefficient	5.12***	5.31***	5.56***	5.82***	-0.35	-0.023	-0.12	-0.12
	(0.44)	(0.59)	(0.44)	(0.64)	(0.34)	(0.34)	(0.34)	(0.34)
Observations	**2784**	**2784**	**2753**	**2753**	**2790**	**2790**	**2759**	**2759**
Quintile Reg. (ITT)								
Coefficient	5.19***	3.06***	5.47***	2.88***	0.086	0.59*	0.041	0.45
	(0.49)	(0.69)	(0.54)	(0.70)	(0.32)	(0.33)	(0.33)	(0.34)
Observations	**2784**	**2784**	**2753**	**2753**	**2790**	**2790**	**2759**	**2759**
2SLS (LATE)								
Coefficient	6.34***	3.76***	6.48***	3.93***	-0.015	0.83	-0.12	0.71
	(0.82)	(1.01)	(0.84)	(1.05)	(0.70)	(0.74)	(0.71)	(0.74)
Observations	**2784**	**2784**	**2753**	**2753**	**2790**	**2790**	**2759**	**2759**

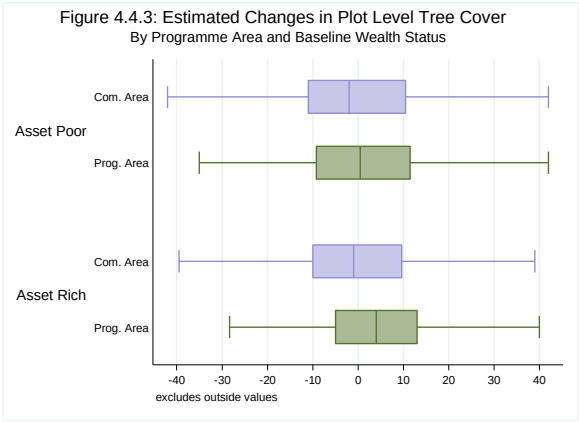
* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level for OLS and 2SLS models; covariates correlated with programme area (p=<0.1) used in all models; IPWRA models also include covariates correlated with outcome (p=<0.1) derived via stepwise regression; VSZ dummies used for fixed effects in all linear models, while exact matching within VSZ enforced for nearest neighbour matching models.

However, as is the case with the Agroforestry Index, our subgroup analyses (see Table A5.2 in Annex 5) revealed considerable heterogeneity in plot-level changes in tree cover since the baseline period. While Figure 4.4.1 reveals a positive overall trend in favour of the programme area, the increases in tree cover were clearly greater in the Tongaren/Likuyani VSZ at around 7%. This was over twice as much of an increase, as compared the Sirisia/Malakisi and Bumula VSZs. While the trend appears positive for the programme area plots in the Kimilili/Ndivisi VSZ, the application of OLS and robust and quantile regression models washes this out. Figure 4.4.4 presents an alternative way of examining the spatial variation in plot-level tree cover changes in the impact study area from 2007 to 2016 by presenting village level average estimates of such change.

Tree cover increased more in the programme area overall, but with significant variation among key subgroups.







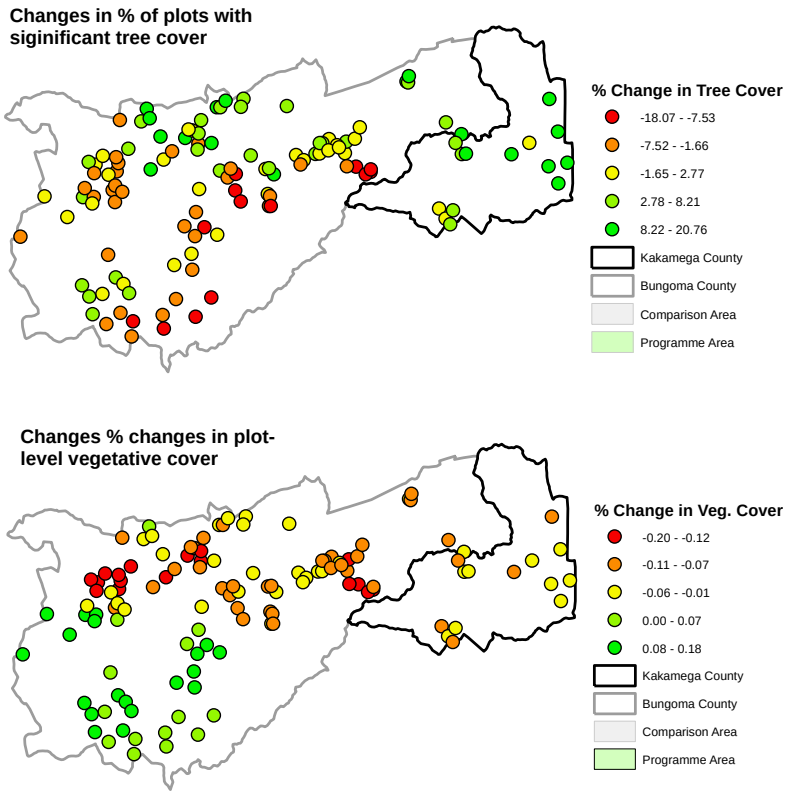
Moreover and as revealed in Table A5.2, while both the plots belonging to households with female and male farmer group members in the programme area experienced positive increases in tree cover vis-à-vis their counterparts in the comparison area, the later did so to a significantly greater extent—4.5% compared to 2%. And the bulk of these positive increases took place in the plots of male headed, as opposed to female headed, households in the programme area, as is revealed by the box plots of Figure 4.4.2 and in Table A5.2. The final significant noteworthy differential effect yielded by our subgroup analyses pertains to household wealth status. As is clear from Figure 4.4.3 and Table A5.2, while positive gains in tree cover were made among relatively poorer households of the programme area, the estimated effect size is only 1.65% (p=0.081). By contrast, this gain is over 4% (p=0.000) for those that were relatively better off in term of asset wealth in the baseline period.

While there is robust evidence that Vi's programme positively impacted changes in tree cover in the programme area, this does not seem to have been the case for general vegetative cover, as the figures presented in the right hand side of Table 4.4.1 reveal. Indeed, the prevalence of general vegetative cover on farm plots was less in 2016 as compared with 2007 in both the programme and comparison areas. However, our subgroup analyses again revealed considerable variation among the VSZs in

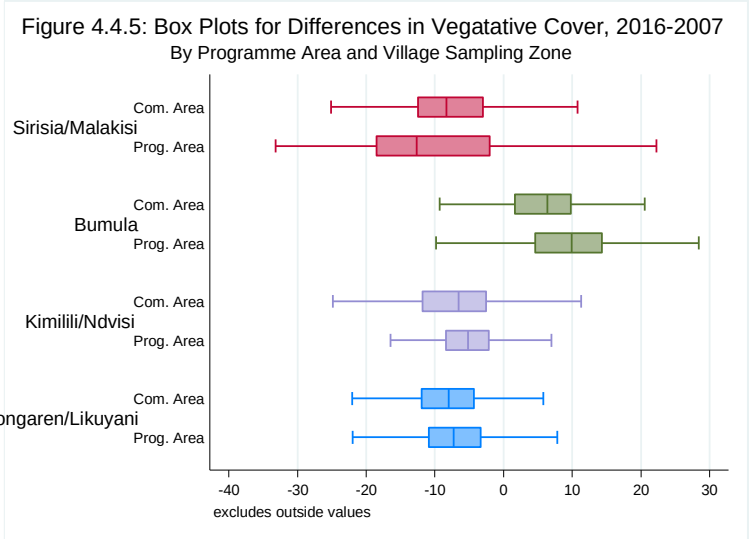
particular. While a general sense of this variation can be seen in Figure 4.4.4, the box plots presented in Figure 4.4.5 reveal this much more overtly. As is clear, Bumula is the only VSZ where plot-level vegetative cover increased in both the programme and comparison areas, and even more so in the former. The trend is in the reverse direction for the Sirisia/Malakisi VSZ, and the differences for the other two are not statistically nor practically significant.

Figure 4.4.4: Village Level Average Estimated Changes in Plot-level Tree and Vegetative Cover, 2007/8 to 2016 (derived from 30m Landsat imagery)

There is little evidence, at least overall, that Vi's programme significantly increased the general prevalence of plot-level vegetative cover.



Changes in fractional vegetative cover vary greatly among the four Village Sampling Zones.



In this subsection remote sensing derived estimates of changes in plot-level tree cover and general vegetative cover were presented. The results for the former complement those of the Agroforestry Index, i.e. that there was considerable uptake of the agroforestry dimensions of Vi's programme. Indeed and overall, plot-level tree cover increased by about 3% among programme area households, while it decreased slightly among those in the comparison area. But this did not happen uniformly. Increases in tree cover were significantly greater in the Tongaren/Likuyani VSZ and among male headed and relatively richer households. Moreover, these results for plot-level tree cover do not transfer over to our complementary fractional vegetative cover measure. The overall difference between this measure in the endline and baseline time periods is marginal, but with widely differing trends among the VSZs.

4.5 Remote Sensing Derived Soil Health Estimates

Table 4.5.1 presents mean and medium values for the soil organic carbon (SOC) and soil erosion prevalence measures, as well as model derived results comparing the plots in the programme and non-programme areas. For SOC, an estimate of 10 is equivalent to 1% organic carbon, 20 2%, and so forth. Estimates below 10 (1% SOC) are generally considered to be low, adversely impacting crop production potential.[5] As indicated, the average values of SOC are, overall, relatively high for both the programme and non-programme areas at around 27 grams per kilogram (2.7%), consistent with the impact study area being located in what is considered part Kenya's high potential area for agricultural production. Nevertheless, the plots in the programme area gained, on average, over one gram of organic carbon per kilogram of soil more than those located in the non-programme area. However, the medium gain is less at 0.88 grams per kilogram of soil. Note that no significant differences between the estimates for all household plots and those specifically focused on food and horticultural crops were observed.

Overall—but with significant site-specific variation—soil organic carbon increased to a greater extent in the programme area vis-à-vis the comparison area, and so too did soil erosion, albeit to a lesser extent.

The figures for soil erosion indicate the percentage of estimated soil erosion prevalence per 30 meter pixel. As shown, soil erosion prevalence is around 40% on average for both the programme and non-programme areas, with it being slightly higher with respect to the former. Note that the estimated differences between the two areas generated through the robust regression and quantile regression models are larger and more statistically significant than the others, revealing that the differences are greater between the two groups for the more typical values in the distribution.

Figure 4.5.1 and Figure 4.5.2 present the distributions of these estimates as box plots. It is clear that the plots in the programme area started off in a better position vis-à-vis those of the comparison area with respect to SOC, while the two areas were about at the same level in estimated levels of soil erosion. There were clear gains in SOC in both areas over the two periods, while the variance in erosion increased considerably. Nevertheless, while the gains in SOC was relatively greater in the programme area overall, there was also a slightly greater overall relative increase in soil erosion as well.

[5] These figures can also be converted into the percentage of organic matter (which includes elements of the soil other than carbon, such as hydrogen, nitrogen, and oxygen) using a factor of 1.72. Hence, 20 grams of SOC per kilogram is equivalent to approximately 3.44% of soil organic matter (20/10 x 1.72).

TABLE 4.5.1: Remote Sensing Derived Land Health Estimates for Main Parcel Plots

	Grams of organic carbon/kg of soil				*% of plot area with significant soil erosion*			
	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.
Raw results								
Sample Mean	26.98	3.31	26.99	3.28	0.39	0.03	0.39	0.03
Program Mean	28.29	3.92	28.26	3.88	0.40	0.04	0.40	0.04
Non-Program Mean	25.64	2.69	25.69	2.65	0.38	0.02	0.38	0.02
Unadjusted dif.	2.65***	1.23***	2.58***	1.23***	0.017***	0.017**	0.019***	0.018**
	(0.36)	(0.26)	(0.37)	(0.27)	(0.0063)	(0.0072)	(0.0064)	(0.0074)
Observations	**2790**	**2790**	**2759**	**2759**	**2790**	**2790**	**2759**	**2759**
Raw results								
Sample Median	23.82	3.41	23.83	3.40	0.36	0.00	0.36	0.00
Program Median	25.23	3.86	25.05	3.82	0.37	0.02	0.37	0.02
Non-Program Median	23.20	2.97	23.20	2.93	0.34	-0.02	0.34	-0.02
Unadjusted dif.	2.04***	0.91***	1.86***	0.89***	0.036***	0.037***	0.039***	0.040***
	(0.66)	(0.28)	(0.65)	(0.30)	(0.0097)	(0.011)	(0.0099)	(0.011)
Observations	**2790**	**2790**	**2759**	**2759**	**2790**	**2790**	**2759**	**2759**
OLS (ITT)								
Coefficient	1.56***	1.15***	1.57***	1.15***	0.016*	0.015	0.017**	0.015
	(0.46)	(0.43)	(0.48)	(0.44)	(0.0084)	(0.0096)	(0.0085)	(0.0097)
Observations	**2790**	**2790**	**2759**	**2759**	**2790**	**2790**	**2759**	**2759**
IPWRA (ITT)								
Coefficient	1.50***	1.48***	1.46***	1.46***	0.015***	0.015***	0.016***	0.016***
	(0.24)	(0.24)	(0.25)	(0.25)	(0.0044)	(0.0044)	(0.0045)	(0.0045)
Observations	**2790**	**2790**	**2759**	**2759**	**2790**	**2790**	**2759**	**2759**
Robust Reg. (ITT)								
Coefficient	1.49***	1.04***	1.51***	1.05***	0.024***	0.021***	0.025***	0.021***
	(0.24)	(0.24)	(0.24)	(0.25)	(0.0043)	(0.0053)	(0.0044)	(0.0054)
Observations	**2790**	**2790**	**2759**	**2759**	**2790**	**2790**	**2759**	**2759**
Nearest Neighbour 1 to 1 Matching (ITT)								
Coefficient	1.20***	1.41***	1.36***	1.39***	0.015***	0.015**	0.015***	0.015**
	(0.29)	(0.29)	(0.32)	(0.30)	(0.0053)	(0.0059)	(0.0054)	(0.0060)
Observations	**2790**	**2790**	**2759**	**2759**	**2790**	**2790**	**2759**	**2759**
Quintile Reg. (ITT)								
Coefficient	1.06***	0.88***	1.19***	0.95***	0.028***	0.029***	0.028***	0.028***
	(0.31)	(0.27)	(0.31)	(0.29)	(0.0054)	(0.0064)	(0.0057)	(0.0062)
Observations	**2790**	**2790**	**2759**	**2759**	**2790**	**2790**	**2759**	**2759**
2SLS (LATE)								
Coefficient	1.97***	1.47***	1.99***	1.47***	0.020*	0.019	0.022**	0.019
	(0.59)	(0.55)	(0.61)	(0.57)	(0.011)	(0.012)	(0.011)	(0.012)
Observations	**2790**	**2790**	**2759**	**2759**	**2790**	**2790**	**2759**	**2759**

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level for OLS and 2SLS models; covariates correlated with programme area (p=<0.1) used in all models; IPWRA models also include covariates correlated with outcome (p=<0.1) derived via stepwise regression; VSZ dummies used for fixed effects in all linear models, while exact matching within VSZ enforced for nearest neighbour matching models.

Figure 4.5.1: Box Plots Comparing Overall Changes in SOC
By Programme Area

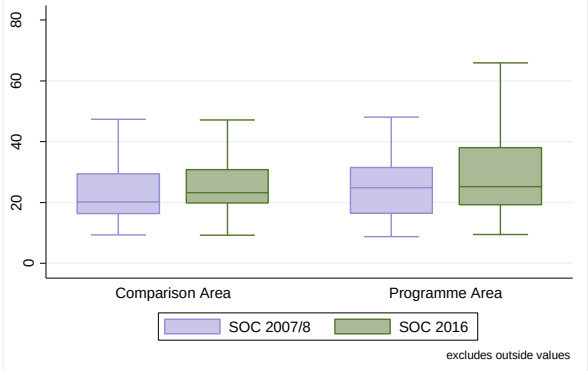


Figure 4.5.2: Box Plots Comparing Changes in Soil Erosion Prevalence
By Programme Area

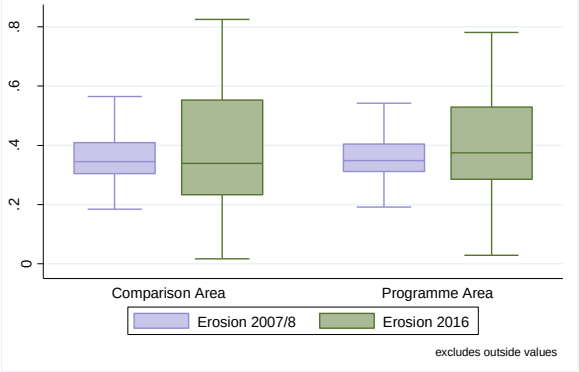
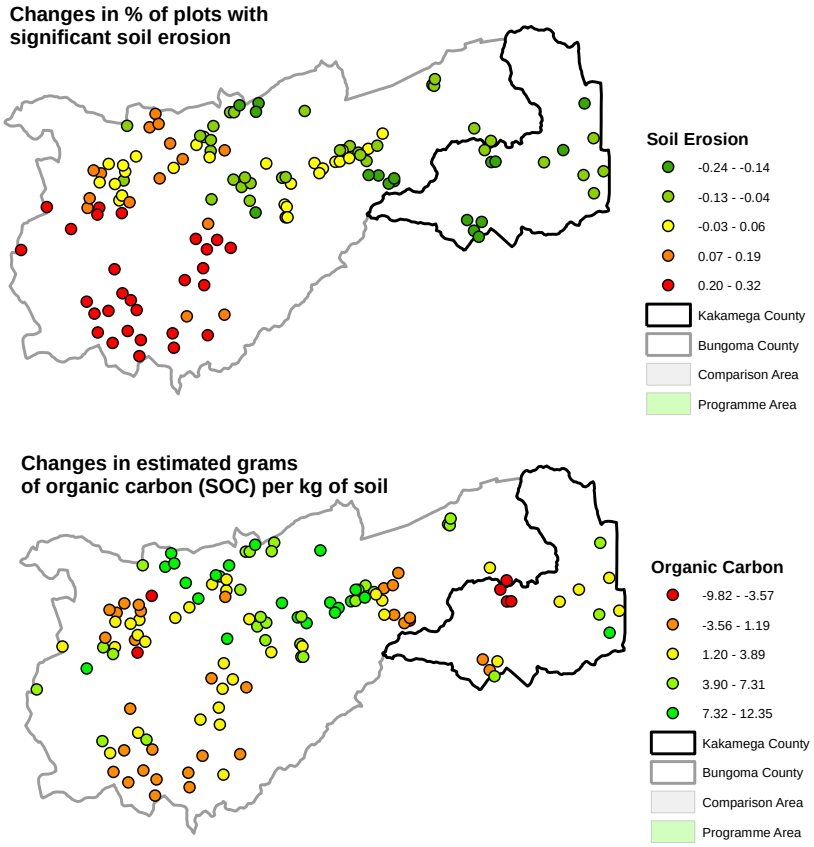


Figure 4.5.3: Village Level Average Estimated Changes in Soil Erosion Prevalence & Soil Organic Carbon, 2007/8 to 2016 (derived from 30m Landsat imagery)

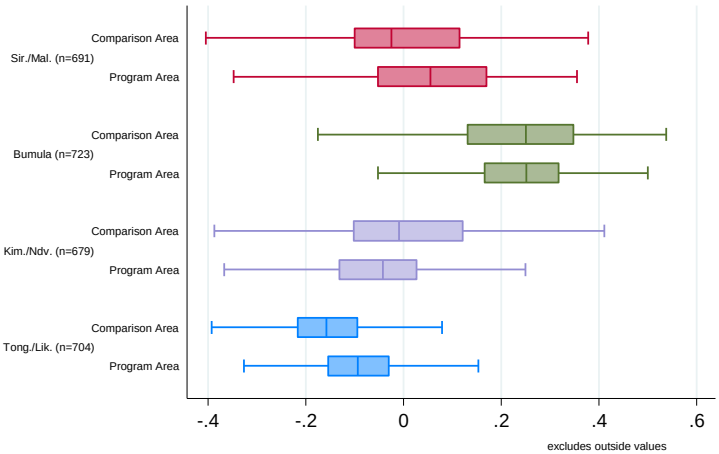


The average and median values presented in Table 4.5.1 mask the site specific changes in both SOC and soil erosion estimated to have taken place from 2007/08 to 2016. Figure 4.5.3 visually presents this spatial variation in village average changes in soil erosion and SOC between the two time periods, respectively. It is clear that both the programme and comparison areas in the western side of the impact study area experienced significant increases in soil erosion, while their counterparts in the east experienced significant reductions. For SOC the spatial pattern is less stark, but there is still considerable variability among the villages in both the intervention and comparison areas.

Figure 4.5.4, combined with results of the sub-group analysis presented in Annex 5, Table A5.3, allows us to explore this spatial variation with respect to estimated soil erosion prevalence further. In the Sirisia/Malakasi VSZ, the prevalence of soil erosion increased more in the programme area than in the comparison area, while in Bumula it increased similarly in both areas. In Kimilili/Ndivisi, the medium rate of soil erosion was about neutral for the comparison area and dropped for most households in the programme area. Finally, while most households experienced a decline in soil erosion in the Tongaren/Likuyani VSZ, the rate was significantly greater in the comparison area.

Relative changes between the programme and comparison area varied considerably among the Village Sampling Zones.

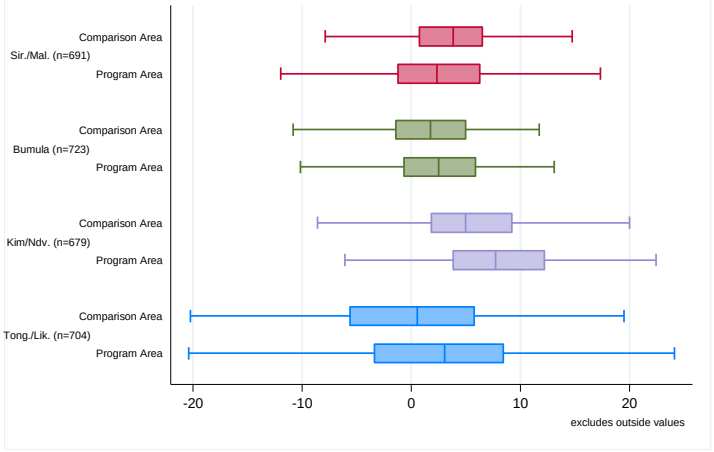
Figure 4.5.4: Changes in Estimated % of Plot-level Soil Erosion
By Program Area & Village Sampling Zone



We found no other sub-group specific differences between the programme and comparison areas for our soil erosion measures, except for whether the respondent holds an official position (see Table A5.3). Interestingly, those who do not hold an official position in the programme area experienced relatively greater rates of soil erosion, as compared with their counterparts in the comparison area. Those with official positions in the programme and comparison area were about on par with respect to changes in soil erosion prevalence.

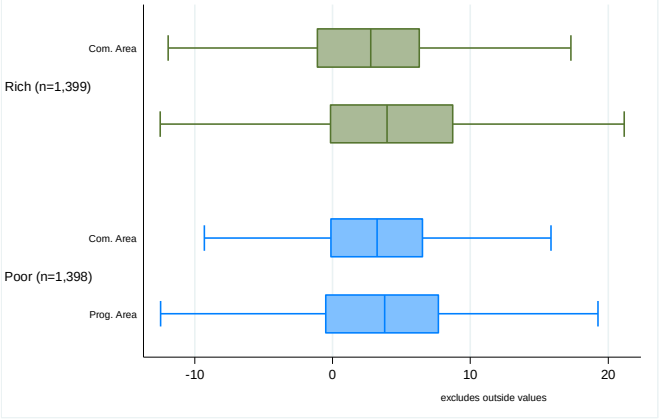
Figure 4.5.5 presents box plots by VSZ, but this time for estimated changes in SOC between 2007/8 and 2016. It is clear that, while most plots experience gains in SOC over the time period, there were only relatively greater gains in the two VSZs located in the eastern part of the impact study area. Moreover, the gains were slightly less in the programme area in the Sirisia/Malakisi VSZ, as compared with the comparison villages in this same VSZ. However, these differences are not independently statistically significant.

Figure 4.5.5: Box Plots for Estimated Changes in Organic Carbon/kg Soil
By Program Area & Village Sampling Zone



Generally speaking, relatively richer households in the programme area were more likely to experience greater rates of SOC accumulation.

Figure 4.5.6: Box Plots for Estimated Changes in Organic Carbon/kg Soil
By Program Area & Baseline Wealth Group



Our further sub-group analysis, presented again in Table A5.3, only revealed one additional sub-group specific difference: those who were richer in asset wealth at baseline experienced relatively greater increases in SOC vis-à-vis the non-programme area as compared with their more baseline asset poor counterparts. Figure 4.5.4 presents this visually as box plots.

In this subsection we examined the extent to which estimates of SOC and soil erosion prevalence differ between the programme and non-programme areas, as well as how these estimates have changed since the baseline period. While plots in the programme area were already better off, overall, in terms of SOC at baseline, they also experienced relatively greater gains overtime. Yet, the prevalence of soil erosion also increased at slightly higher rate in the programme area overall as well.

Moreover, it is critical to acknowledge that there is considerable variation in these estimates and trends among the VSZs. Relative improvements in SOC over the comparison area only took place in the two VSZs in the eastern part of the impact study area, while two of the VSZs—one in the east and one in the west—actually experienced relatively greater increases in soil erosion vis-à-vis the comparison areas. Only one VSZ, Kimilili/Ndivisi, experienced a relative decrease. Moreover, it was only those without official positions in their respective communities in the programme areas that experienced this relatively greater rate of soil erosion, and the baseline rich in these areas were more likely to experience gains in SOC. We therefore conclude that evidence of programme impact on these two soil health measures is mixed.

4.6 Tree Product Sales and Firewood Access

4.6.1 Tree Product Sales

One hypothesized pathway for agroforestry to generate longer term socio-economic impacts is by directly generating income through the sale of agroforestry products, such as timber, fuelwood, and fruits. Our theory of change for Vi's programme therefore postulates that we should see higher earnings from the sale of such products in the programme area in general and among Vi affiliated households in particular vis-à-vis those in the comparison area. Hence, as the enumerators visited each plot of the household's main parcel, they were asked to observe if there were any trees or shrubs within or along the boundaries of the plot and, if so, whether they had generated any products over the last 12 months, such as fodder, timber,

fuelwood, and/or fruits, that were sold for cash. While we recognize the potential for recall bias, they were also asked the same questions about the plot with respect to the baseline period.

TABLE 4.6.1: Tree Product Sales

	Sale of AF Products Ksh (2016)	Sale of AF over Ksh 1000 (2016)	Sale of AF Products (log) (2016)	Sale of AF Products Ksh (difference)	Sale of AF over Ksh 1000 (difference)	Sale of AF Products (log) (difference)
Raw results						
Sample Mean	6733.12	0.30	3.01	4787.21	0.27	11.48
Program Mean	8647.15	0.36	3.60	6347.68	0.33	11.50
Non-Program Mean	4784.57	0.25	2.42	3198.60	0.22	11.46
Unadjusted dif.	3862.6***	0.11***	1.18***	3149.1***	0.10***	0.040***
	(836.6)	(0.017)	(0.16)	(775.1)	(0.017)	(0.013)
Observations	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**
Raw results						
Sample Median	0.00	0.00	0.00	0.00	0.00	11.45
Program Median	0.00	0.00	0.00	0.00	0.00	11.45
Non-Program Median	0.00	0.00	0.00	0.00	0.00	11.45
Unadjusted dif.	0.00	0.00	0.00	0.00	0.00	0.00
Observations	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**
OLS (ITT)						
Coefficient	3428.0***	0.34***	1.06***	3384.9***	0.36***	0.043***
	(886.8)	(0.061)	(0.18)	(858.2)	(0.062)	(0.014)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
IPWRA (ITT)						
Coefficient	No convergence	0.098***	1.03***	No convergence	0.10***	0.041***
		(0.016)	(0.15)		(0.016)	(0.013)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
Nearest Neighbor 1 to 1 Matching (ITT)						
Coefficient	3493.9***	0.098***	0.97***	3683.8***	0.12***	0.044***
	(746.1)	(0.020)	(0.18)	(764.4)	(0.020)	(0.016)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
2SLS (LATE)						
Coefficient	4344.4***	0.43***	1.34***	4290.2***	0.45***	0.055***
	(1118.8)	(0.077)	(0.23)	(1082.0)	(0.078)	(0.018)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level for OLS and 2SLS models; covariates correlated with programme area (p=<0.1) used in all models; IPWRA models also include covariates correlated with outcome (p=<0.1) derived via stepwise regression; VSZ dummies used for fixed effects in all linear models, while exact matching within VSZ enforced for nearest neighbour matching models.

Table 4.6.1 presents the results comparing the programme and non-programme areas. The three single differenced models include the recalled baseline estimates for the reported cash value of what agroforestry products were reported as being sold. There are, again, several noteworthy observations. First, there are significant differences in favour of programme area in terms of income from reported agroforestry product sales, both in general and in terms of changes from the baseline. However, the second observation is that most households in *both* the programme and non-programme areas reported no sales at all; the median is 0 for both groups. We further created two binary variables, one to indicate that the household had made a significant sale of one or more agro-forestry products over the last 12 months, i.e. over Ksh 1,000 (≈USD $10), and other for a difference in earnings of over Ksh 1,000 from the baseline period. As is clear from Table 4.6.1, only just over one-third of households in the programme area scored positively on the former indicator, as

compared with about one-fourth in the comparison area. The results are similar for the differenced binary variable.

While average income from the sale of agroforestry products in certainly higher in the programme area, there is considerable variation among households and the VSZs, with most households not reporting any sales at all.

Table A5.4 in Annex 5 presents the results of our subgroup analyses. The most significant differences are in relation to the interviewed respondents residing in the different VSZs. In particular, there are much larger differences between the programme and non-programme areas in the Sirisia/Malkisi and Bumula VSZs and to a lesser extent in the Tongaren/Likuyani VSZ. The overall difference between the programme and non-programme areas and the VSZ specific differences are clear in the box plots presented in Figure 4.6.1, which shows reported changes in income from the sale of agroforestry products between the two time periods. Note that outside values have been excluded, which makes it appear that there is a significant difference between the programme and non-programme areas even in the Kimilili/Ndivisi VSZ. However, these outside values were included in the analyses informing Table A5.4, so the apparent difference is statistically insignificant. Unfortunately, due to the fact that the majority of households reported no sales of agroforestry products, both our robust and quantile regression models failed to converge, as did two of our IPWRA models.

Figure 4.6.1: Box Plots for Changes in Income from Agroforestry Products
By Program Area & Village Sampling Zone

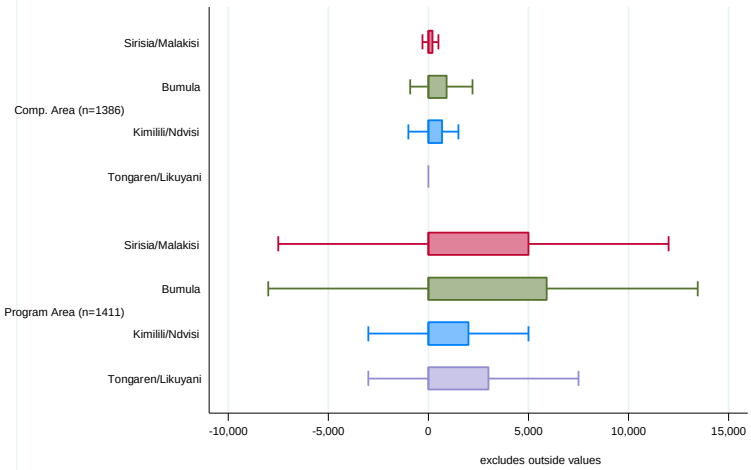
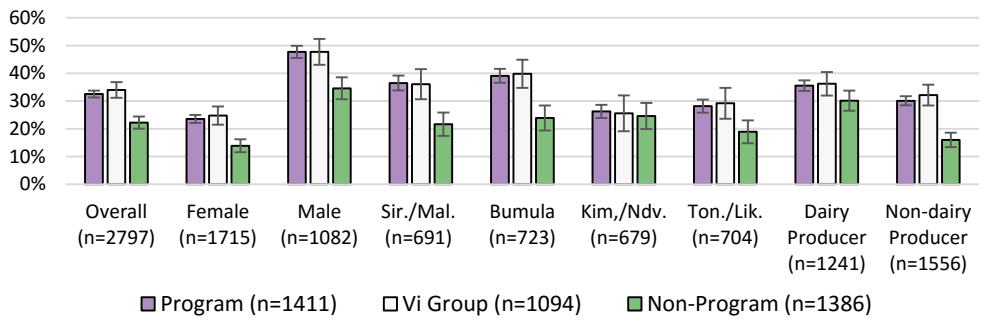


Figure 4.6.2 presents the results of complementary analysis of respondents reporting over Ksh 1,000 additional income from the sale of agroforestry products between the two time periods, i.e. the second binary indicator in Table 4.6.1. While the differences between female and male respondents is clear, the relative difference vis-à-vis their counterparts in the non-programme areas is statistically insignificant. Note also that the differences between the programme area and Vi groups on the one hand and non-programme area on the other are significant, save for the Kimili/Ndivisi VSZ. It is interesting also that dairy producers in the non-programme area are nearly on par with those in the programme area, while the reverse is the case among non-dairy producers. This is consistent with the findings for the Agroforestry Index presented above; the adoption of Vi's general agroforestry programme took place to a greater extent among non-dairy producers.

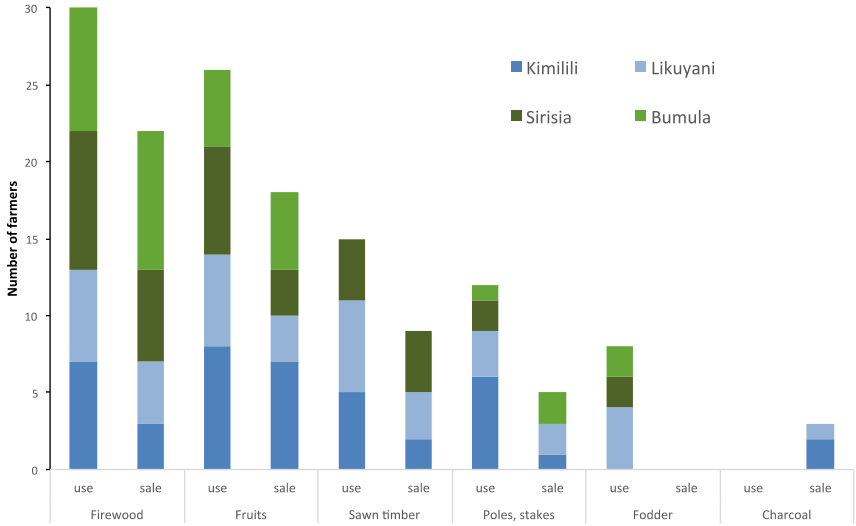
Figure 4.6.2: Reported Increases in Sales of Agroforestry Products (> Ksh1000): 2007 versus 2016 by Gender & Village Sampline Zone



Our in-depth qualitative work also explored both the use and sale of particular agroforestry products, and those products presented in Figure 4.6.3 were elicited. Most of those interviewed source firewood from their own farms, and it is the product most sold across the study area, with the highest numbers in Bumula and lowest in Kimilili. The most dominant firewood tree species found is *Sesbania sesban* (particularly in Bumula), followed by *Markamia lutea* and *Grevillea robusta*. A quarter of farmers visited were harvesting the former for firewood for sale, with average annual reported income ranging from USD $10 to $120 (the latter was reported by a women farmer in Bumula harvesting three tons of *Calliandra* and *Sesbania* every year). Five farmers reported earnings of over USD $50 per year from *Sesbania*.

Firewood was revealed through our qualitative work to be the most popular agroforestry product for both use and sale, followed by fruits.

Figure 4.6.3: Use & Sales of AF Products Revealed by Qualitative Interviews



Fruit trees were found on 65% of the farms, but relatively fewer farmers reported selling fruit in Likuyani and Sirisia. Most of the 18 farmers reported modest income from the sales of fruit, i.e. less than USD $10. However, four farmers reported earning over USD $100. Moreover, fifteen of the 40 interviewed farmers reported accessing timber from their farms, with Eucalyptus being the dominate species. Only nine reported timber sales, with most selling only one or two trees over the last five years,

earning approximately USD $15 to $40 per tree. Three farmers had woodlots, however, and reported earnings of over USD $250 over this same time period. Less than a quarter of farmers reported using tree fodder, and this is only for their own livestock feeding needs. Finally, it is worth noting that three farmers in Kimilili and Likuyani produced charcoal for sales purposes only.

4.6.2 Firewood Cash Value and Collection Time

Recall that another expected intermediary outcome associated with the increased uptake of agroforestry is increased access fuelwood, given that it can be readily obtained from the household in question's farm. This, in turn, is expected to reduce the amount to time and effort from having to collect such wood from further afield, including local forests. Given that nearly all households in the impact study area are dependent on fuelwood (>99%), coupled with the gender-based division of labour with respect to its collection, we assume that reducing the amount of time and effort spent collecting it would positively benefit women as well. It may further lead to increased income and other related benefits, as their time is freed to pursue other livelihood pursuits.

More trees on farm would imply more firewood being accessed on farm, followed by a reduction in time and effort collecting it.

To capture data on the amount of firewood accessed on farm and the time spent collecting it, the respondents were first asked whether they had used any firewood for cooking, heating, or any other purpose during the previous month. They were then asked where they sourced it from, including the primary source, followed by (a) the number of times they collected it over the past month; (b) the number of hours spent collecting it on each occasion; and (c) how much what was collected would have cost if it were purchased from the local market. Again, while recognizing the inherent propensity for significant recall bias, the respondents were also ask to recall similar information for the average month in 2007. Through these data, we constructed several values pertaining to the estimated cash value of firewood collected on farm and time spent collecting household firewood in a given month. Table 4.6.2 presents the results of a comparison between households of the programme and non-programme areas.

There are, again, some noteworthy results. First, the estimated cash value of firewood collected on farm and increases in this cash value over the two time periods is, overall, greater for the programme area. While most of the robust and quantile regression models yielded statistically significant results, they indicate that influential observations are driving the size of the average differences to a significant extent. Nevertheless, the median gain in firewood cash value is close to the average gain.

Consistent with the Agroforestry Index, firewood availability on farm seems to have increased more in the programme area's western side.

Our sub-group analyses for these variables (see Table A5.5) reveal, again, differences among the VSZs, as evident in Figure 4.6.3. Indeed, the effect estimates for the Sirisia/Malakisi and Bumula VSZs are not only the ones that are independently statistically significant ($p<0.01$) but also are much larger than the other two VSZs. However, with the application of robust regression, Bumula is the only VSZ with a relatively large and independently statistically significant effect size (Ksh 170; p=0.001). The other noteworthy sub-group specific difference is, again, in relation to dairy versus non-dairy farmers, with the latter in the programme area reporting relatively more firewood being collected from their farms (Ksh 201; p=001). The application of robust regression does reduce the effect estimate for this group to Ksh 74, but it still remains independently statistically significant (p=0.035).

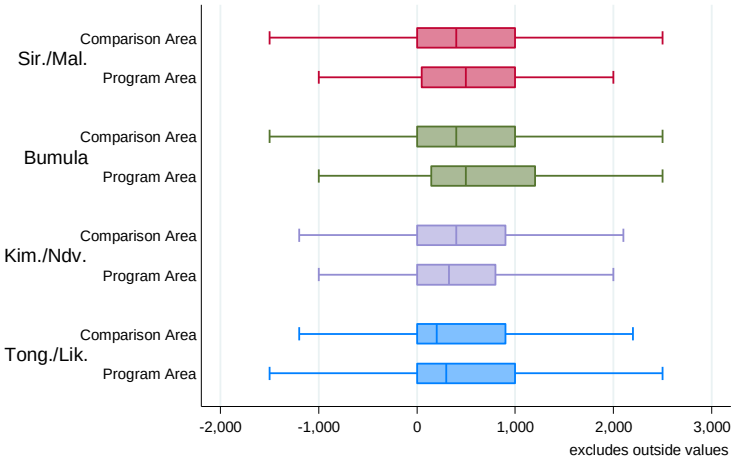
TABLE 4.6.2: Firewood Cash Value and Collection Time

	Cash Value of Firewood from Farm				*Hours in Month Collecting Firewood*			
	2016	***2016 log***	***Dif.***	***Dif.(log)***	***2016***	***2016 log***	***Dif.***	***Dif.(log)***
Raw results								
Sample Mean	1202.60	5.46	574.16	0.96	10.13	2.04	-1.95	-0.10
Program Mean	1280.91	5.81	640.47	1.18	9.43	1.97	-2.44	-0.15
Non-Program Mean	1122.88	5.11	506.65	0.74	10.84	2.10	-1.44	-0.05
Unadjusted dif.	158.0***	0.70***	133.8***	0.44***	-1.41***	-0.13***	-0.99*	-0.097***
	(43.4)	(0.12)	(35.8)	(0.12)	(0.44)	(0.031)	(0.55)	(0.031)
Observations	**2797**	**2797**	**2797**	**2797**	**2784**	**2784**	**2739**	**2739**
Raw results								
Sample Median	1000.00	6.91	400.00	0.47	6.00	1.95	0.00	0.00
Program Median	1000.00	6.91	500.00	0.51	6.00	1.95	0.00	0.00
Non-Program Median	1000.00	6.91	400.00	0.41	6.00	1.95	0.00	0.00
Unadjusted dif.	0	0	100***	0.11***	0	0	0	0
	(34.3)	(0.035)	(34.2)	(0.037)	(0.24)	(0.070)	(0.034)	(0.0082)
Observations	**2797**	**2797**	**2797**	**2797**	**2784**	**2784**	**2739**	**2739**
OLS (ITT)								
Coefficient	143.8***	0.54***	138.4***	0.39***	-1.47***	-0.12***	-0.87	-0.087**
	(40.3)	(0.12)	(39.7)	(0.13)	(0.42)	(0.029)	(0.60)	(0.034)
Observations	**2790**	**2790**	**2790**	**2790**	**2732**	**2732**	**2732**	**2732**
IPWRA (ITT)								
Coefficient	134.1***	0.50***	132.3***	0.38***	-1.35***	-0.11***	-0.69	-0.079**
	(35.6)	(0.099)	(36.0)	(0.12)	(0.40)	(0.027)	(0.56)	(0.032)
Observations	**2790**	**2790**	**2790**	**2790**	**2732**	**2732**	**2732**	**2732**
Robust Reg. (ITT)								
Coefficient	50.9*	-0.0097	63.5**	0.039*	-0.94***	-0.12***	-0.41*	-0.077***
	(26.1)	(0.017)	(27.2)	(0.021)	(0.20)	(0.026)	(0.22)	(0.028)
Observations	**2790**	**2790**	**2790**	**2790**	**2732**	**2732**	**2732**	**2732**
Nearest Neighbour 1 to 1 Matching (ITT)								
Coefficient	120.0***	0.47***	141.7***	0.40***	-0.99**	-0.095***	-0.64	-0.070*
	(40.5)	(0.13)	(42.3)	(0.15)	(0.45)	(0.032)	(0.67)	(0.040)
Observations	**2790**	**2790**	**2790**	**2790**	**2732**	**2732**	**2732**	**2732**
Quintile Reg. (ITT)								
Coefficient	98.1***	0.084	122.9***	0.083**	-1.06***	-0.10***	0	0
	(32.4)	(0.20)	(32.9)	(0.034)	(0.34)	(0.034)	(0.070)	(0.011)
Observations	**2790**	**2790**	**2790**	**2790**	**2732**	**2732**	**2732**	**2732**
2SLS (LATE)								
Coefficient	182.2***	0.68***	175.4***	0.49***	-1.87***	-0.15***	-1.43*	-0.13***
	(50.8)	(0.15)	(50.1)	(0.16)	(0.52)	(0.037)	(0.76)	(0.044)
Observations	**2790**	**2790**	**2790**	**2790**	**2732**	**2732**	**2732**	**2732**

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level for OLS and 2SLS models; covariates correlated with programme area (p=<0.1) used in all models; IPWRA models also include covariates correlated with outcome (p=<0.1) derived via stepwise regression; VSZ dummies used for fixed effects in all linear models, while exact matching within VSZ enforced for nearest neighbour matching models.

If the cash value of firewood collected from the farm is a good proxy for its quantity produced on farm, we would expect to see a corresponding decrease in the amount of hours spent collecting it. The overall results of our analyses are presented in the right hand side of Table 4.6.2. In general, we do see significant differences in favour of the programme area. However, only 44% and 41% of households reported a reduction in the programme and comparison areas, respectively, and with a raw median value for both groups of 0. Moreover, only the robust regression and 2SLS estimates are statistically significant, and this is only at the 10% level. The greater significance of the logged measure of this differed variable reveals that influential observations are likely influencing the results of the statistical tests.

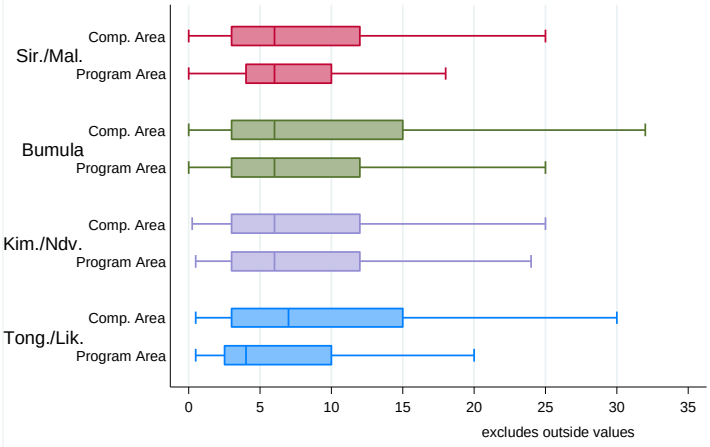
Figure 4.6.3: Cash Value (Ksh) of Fuel Wood Collected from Own Farm
By Programme Area and Zone



Moreover, the results of our sub-group analyses (Table A5.4) again reveals that there are differences among the VSZs, with the Tongaren/Likuyani respondents from the programme area reporting both less time and a greater drop in time spent collecting firewood. This variation is reflected in the box plots presented in 4.6.4.

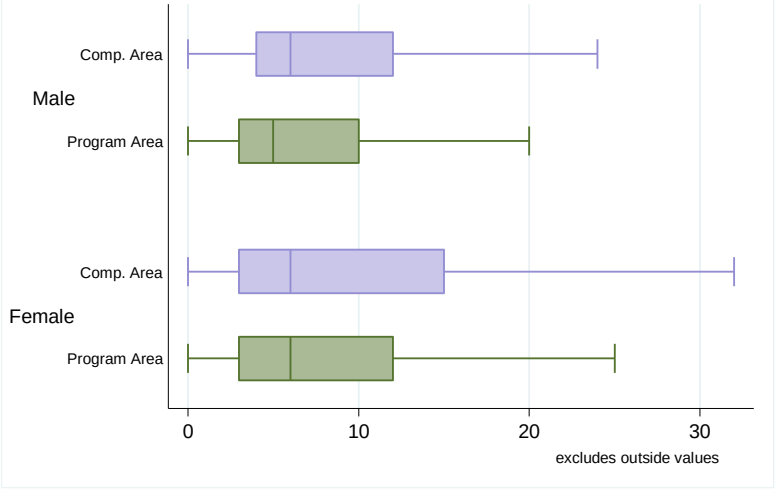
There is variation in the effect estimates for the cash value of firewood and firewood collection time measures among the village sampling zones. Moreover, men from the programme area were more likely to report a drop in firewood collection time than women.

Figure 4.6.4: Reported Hours Spent Collecting Firewood Per Month (2016)
By Program Area & Village Sampling Zone



The other noteworthy subgroup related difference pertains to male versus female respondents. Men from the programme area were more likely to report a decrease in hours spent collecting firewood vis-à-vis their male counterparts in the comparison area, but this is not the case for female respondents (Figure 4.6.5). This is perplexing considering that the latter are the ones who traditionally collect it.

Figure 4.6.5: Reported Hours Spent Collecting Fuelwood Per Month (2016)
By Program Area and Gender



In this Subsection, we have seen that the trends in changes in income from the sale of agroforestry products and the estimated cash value of fuelwood obtained on farm mirror—at least to an extent—those found for the Agroforestry Index. In particular, while both did increase significantly more, on average, among respondents in the programme area, this was highly variable, particularly among the VSZs. Dairy farmers in the programme area do not seem to have been shifted as much as non-dairy farmers in relation to these measures as well. While many of the results are statistically significant, evidence that the impact of Vi's programme in reducing time spent collecting firewood is less robust and clear-cut, particularly given that it is only male respondents from the programme area who were more likely to report drop in firewood collection time and not their female counterparts.

4.7 Tree Fodder Use and Milk Yields

Vi's programme was expected to benefit dairy farmers, specifically through the uptake of tree/shrub fodder, which was expected to, in turn, boost milk yields.

One of our sub-group effect hypotheses is that dairy farmers were impacted by Vi's programme more than non-dairy famers, given the income generation potential of dairy production interacted with the previously evidenced efficacy of particular species tree/shrub fodder. However, for our Agroforestry Index and the tree product sales and the cash value of fuelwood measures, we have seen that they appear to have gained little from Vi's programme. In this subsection we narrow in specifically on these dairy farmers to assess whether there is evidence that they took up the component Vi's programme for which they were exclusively targeted—the production and use of tree/shrub fodder—and whether milk yields were boosted.

In the questionnaire's livestock ownership module, the respondents were first asked whether their household owns any livestock and, if so, the particular types (including improved and local dairy cows and improved and local dairy goats), followed by specific questions about each. For dairy animals, this included questions about (a) the average yield of milk produced per animal during the milking season, both at present and in 2007; and (b) whether commercial feed and/or tree fodder is (or was in 2007) fed to the animals. From these data, we constructed the outcome measures presented in Table 4.7.1.

TABLE 4.7.1: Tree Fodder Use and Milk Yields among Dairy Producers

	Tree fodder use 2016	*Tree fodder use dif.*	*Avg. 2016 milk ltr./day*	*Avg. dif. ltr./day*	*Avg. dif. ltr./day (log)*	*% change in milk yields*	*+ change in milk yields*	*Self-reported increase*
Raw results								
Sample Mean	0.32	0.18	3.50	0.27	0.10	0.23	0.47	0.48
Program Mean	0.42	0.27	3.63	0.39	0.14	0.28	0.50	0.52
Non-Program Mean	0.21	0.10	3.37	0.17	0.07	0.18	0.44	0.44
Unadjusted dif.	0.21***	0.17***	0.26*	0.22**	0.070**	0.099**	0.061*	0.077**
	(0.030)	(0.027)	(0.14)	(0.10)	(0.031)	(0.044)	(0.033)	(0.033)
Observations	**932**	**932**	**932**	**932**	**931**	**932**	**932**	**928**
Raw results								
Sample Median	0.00	0.00	3.00	0.00	0.00	0.00	0.00	0.00
Program Median	0.00	0.00	3.00	0.00	0.00	0.00	0.00	1.00
Non-Program Median	0.00	0.00	3.00	0.00	0.00	0.00	0.00	0.00
Unadjusted dif.	--	--	0	0	0	0	--	--
			(0.12)	(0.082)	(0.033)	(0.037)		
Observations	**932**	**932**	**932**	**932**	**931**	**932**	**932**	**928**
OLS (ITT)								
Coefficient	0.63***	0.73***	0.30***	0.27***	0.082**	0.12**	0.20**	0.20**
	(0.11)	(0.11)	(0.10)	(0.10)	(0.032)	(0.046)	(0.090)	(0.086)
Observations	**929**	**929**	**929**	**929**	**928**	**929**	**929**	**925**
IPWRA (ITT)								
Coefficient	No conveg.	0.16***	0.30***	0.26**	0.068**	0.11**	0.073**	0.075**
		(0.028)	(0.095)	(0.10)	(0.031)	(0.044)	(0.033)	(0.033)
Observations	**929**	**929**	**929**	**929**	**928**	**929**	**929**	**925**
Robust Reg. (ITT)								
Coefficient	--	--	0.22***	0.20**	0.066**	0.064**	--	--
			(0.081)	(0.080)	(0.029)	(0.032)		
Observations			**929**	**929**	**928**	**929**		
Nearest Neighbour 1 to 1 Matching 1 to 1 (ITT)								
Coefficient	0.19***	0.18***	0.24*	0.21*	0.079**	0.13**	0.065*	0.093**
	(0.032)	(0.033)	(0.12)	(0.12)	(0.036)	(0.052)	(0.039)	(0.038)
Observations	**929**	**929**	**929**	**929**	**928**	**929**	**929**	**925**
Quintile Reg. (ITT)								
Coefficient	--	--	0.23**	0.13	0.041	0.045	--	--
			(0.090)	(0.088)	(0.033)	(0.038)		
Observations			**929**	**929**	**928**	**929**		
2SLS (LATE)								
Coefficient	0.77***	0.85***	0.36***	0.33***	0.10***	0.14***	0.24**	0.25**
	(0.13)	(0.14)	(0.12)	(0.12)	(0.039)	(0.056)	(0.11)	(0.10)
Observations	**929**	**929**	**929**	**929**	**928**	**929**	**929**	**925**

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level for OLS and 2SLS models; covariates correlated with programme area (p=<0.1) used in all models; IPWRA models also include covariates correlated with outcome (p=<0.1) derived via stepwise regression; VSZ dummies used for fixed effects in all linear models, while exact matching within VSZ enforced for nearest neighbour matching models.

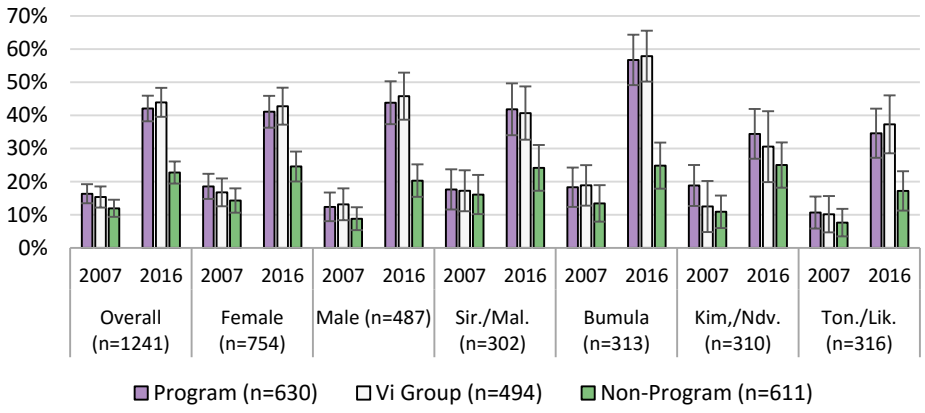
For the milk yield measures, we restricted our analysis to cows only, given that there are only seven respondents who reported owning improved dairy goats in 2007. Our rationale for doing so is also related to another variable construction procedure we followed. Rather than taking the average of all the reported milk yields across improved and local cows by time period and then differencing these averages, we first differenced by cow type and then took the averages across these differences. If a household only owned one cow type in both time periods, then the average pertains to only this type. Moreover, if a household had only local cows in 2007 and then added one or more improved cows to its dairy portfolio, only the average changes in milk yield pertaining to the former were assessed. This was to ensure that the estimated changes in milk yield are not a reflection of a household having had, for example, upgraded from one more local cows to one or more improved cows.

The first two columns in Table 4.7.1 present single and double difference measures for reported tree/shrub fodder use for feeding dairy animals, respectively. Approximately, twice as many dairy producers in the programme area make use of dairy fodder, and this increased by 27% against 10% among their counterparts in the comparison area.[6] Figure 4.7.1 further highlights the overall difference and also disaggregates the results by respondent gender and VSZ. Here, we can see that the results are quite similar along gender lines but with clear variation among the VSZs, with Bumula showing the greatest difference between the programme and non-programme areas and Kimilili/Ndivisi—yet again—the least.

There is evidence that Vi's programme led to greater tree/shrub fodder use and increased milk yields in the programme area, but only among some of the participating dairy farmers and with some VSZs benefiting more than others.

For average milk yields and increases in milk yields, there are consistent positive results, despite the fact that the median raw difference in both the programme and comparison areas is zero. Nevertheless, we do see that reported milk yields increased in favour of the programme area by at least 0.2 litres per day or at least 6% (using the more conservative robust regression estimates). A greater percentage of dairy producers in the programme area also qualitatively reported that their milk yields had increased from the baseline period—52% against 44%.

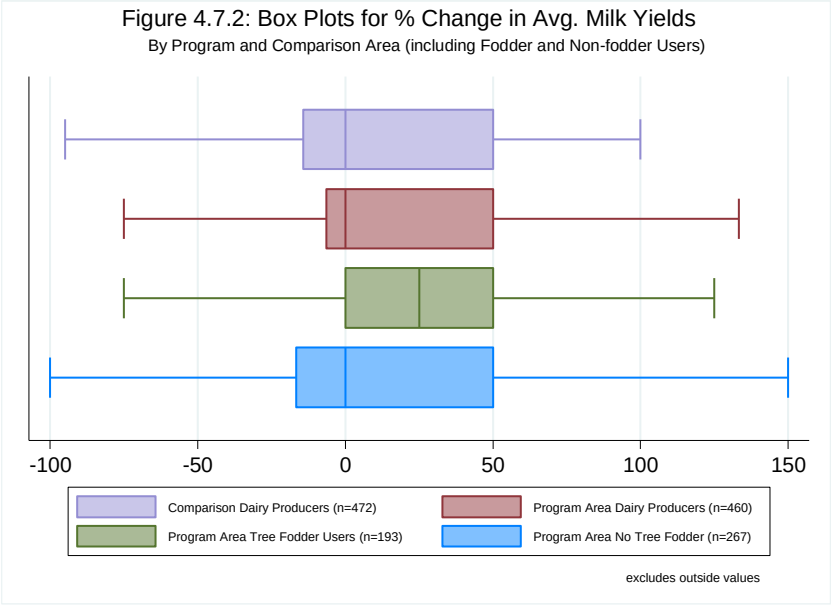
Figure 4.7.1: Tree Fodder Use Among Dairy Producers by Gender & Village Sampling Zone



A relevant question, of course, is: To what extent was the relatively greater increase in milk yields among the dairy farmers in the programme area driven by their relatively greater use of tree/shrub fodder? While we recognize that comparing milk yields between tree/shrub fodder users and non-users would be inconclusive (i.e. there may be one or more 'omitted' variables correlated such differential uptake that could account for the difference), failing to see such a relationship in the data would give us strong grounds to reject this as a hypothesised mechanism. In Figure 4.7.2, we present four box plots for the differenced milk yield measure: the first is for the comparison area and the second for the overall programme area, while the third and fourth are specific to tree/shrub fodder and non-fodder users residing in the latter, respectively. It is clear that changes in milk yields between the comparison area and non-tree/shrub fodder users in the programme area are very similar. The box plot for the tree/shrub fodder users in the programme area clearly standout, with three-quarters of the distribution reporting positive milk yield increases.

[6] Given that the dairy farmers are a sub-set of the overall sample, a specific set of covariates correlated with programme area (at p<.1) specific to this sub-sample was used.

Figure 4.7.2: Box Plots for % Change in Avg. Milk Yields
By Program and Comparison Area (including Fodder and Non-fodder Users)

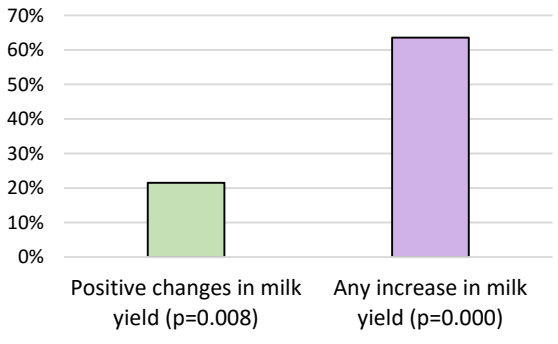


To further explore whether the increased uptake of tree/shrub fodder was a possible mechanism responsible for bolstering milk yields to a relatively greater extent in the programme area, we carried out mediation analysis using Stata's *sem* (structural equation modelling) command. Given that the uptake of tree/shrub fodder was voluntary (i.e. non-random), positive results associated with such analysis would not conclusively evidence that this was certainly a mechanism; as noted above, other unaccounted factors correlated with tree/shrub fodder uptake may be responsible. However, the data would nevertheless support the plausibility of it being a key mechanism, while the absence of such positive results would prompt us to rule it out as being a likely contender.

The data are consistent with tree/fodder use being at least partly responsible for why milk yields were reported as higher, on average, in the programme area.

While we cannot rule out the possibility of 'intermediate variable bias', we attempted to mitigate it by using a stepwise regression model to identify which of our covariates are correlated with tree/shrub fodder uptake at the 10% level or less and then included these in our mediation model. Figure 4.7.3 shows how much of the programme area's effect on both increases in milk yield (continuous variable) and any increase in milk yield (binary variable) is explained through its effect of the uptake of tree/shrub fodder. While the variation in the data shared by both tree/shrub fodder uptake and programme area are significantly associated with positive increases in milk yield (indirect effect=0.05; p=0.004), the former only accounts for (i.e. mediates) about one-fifth of the former's total effect. This is perhaps not surprising, given that there are many other factors that can affect changes in milk yield, including the quantity and regularly of the feeding of the tree/shrub fodder itself. Indeed, it is clear that much more of the

Figure 4.7.3: % of Total Estimated Effect of Programme Area on Increased Milk Yields Mediated by Uptake of Tree/Shub Fodder



total effect of the programme area dummy variable is mediated by tree/shrub fodder uptake for the binary milk yield increase measure (64%), given that all the nuanced variation associated with the continuous measure has been stripped away.

In this subsection we have seen that, while dairy farmers appear to have taken up less of Vi's overall programme, many did engage with its tree/shrub fodder promotion activities. However, the use of such fodder among farmers in the programme area that had been engaged in dairy farming since the baseline period only increased by 17%, with only 42% reporting using it. Nevertheless, there is fairly strong and consistent evidence that Vi's programme supported many dairy farmers in the programme area to bolster milk yields, and it is quite plausible that their uptake of tree/shrub fodder promoted through this same programme was at least partially responsible.

Thus far we have seen that there was a significant, yet variable and somewhat limited, uptake of Vi's programme in the programme area and that similarly variable effects were observed for most of the above intermediary impact measures. We will now examine the extent this programme impacted several more downstream impact measures, including household consumption expenditure, several asset-based indices, the Coping Strategy Index, Minimum Dietary Diversity-Women (MDD-W), months of adequate food provisioning, and education progression and spending.

4.8 Household Consumption Expenditure and Asset Wealth

Measuring household wealth or income in low and middle income countries is not straightforward, particularly in rural areas where respondents tend to be self-employed. Self-reported measures of total income can be unreliable, given that such households are often involved in a wide variety of livelihood pursuits (Morris et al. 2000). However, given that there is a recognised and strong association between household income and consumption, one popular proxy measure used by the World Bank and other international institutions is household consumption expenditure (Deaton and Zaidi 1999). It is through such data that the percentages of households living above and below the poverty line are estimated.

Measuring total household income in low and middle income countries can be difficult, given the variety of livelihood pursuits in which households are engaged. Looking at outflows—i.e. consumption expenditure—is therefore seen as a viable alternative.

To collect such data, several modules were incorporated into the household survey. The respondents were asked, for instance, the types of food their households consumed over the previous seven day period, as well as the particular quantities. These quantities were then converted into monetary values. This was done by asking the respondent how much was paid for each food item or, if the food item was sourced through the household's own production, how much it would have costed if it were purchased from the local market. The respondents were also asked how much they spent on particular non-food items and services from a detailed list, such as soap, toothpaste, and minibus fares, over the past four weeks (regular non-food expenditure). Finally, they were asked about particular 'big ticket' expenditures over the previous 12 months from another pre-defined list, such as school and hospital fees, clothes, and home repair (irregular non-food expenditure).

We then computed the basic per capita measure as follows for each household:

- The weekly cash value of each food item consumed during the past seven days were added together and divided by seven, thereby estimating the daily cash value of food consumed by the household.

- Household expenditure on items from both the regular monthly non-food expenditure list and annual non-food expenditure list were added together and divided by 30.42 and 365, respectively, thereby estimating the household's average daily expenditure on regular and irregular non-food items.
- The daily consumption expenditure estimated for food and the regular and irregular non-food items were then added together and converted into US dollars, while adjusting for purchase power parity (PPP).[7]
- Finally, to derive each household's per capita consumption expenditure, its PPP adjusted dollar value was divided by the number of its members (household size), with another adjustment made for assumed lower consumption among children and economies of scale.[8]

Table 4.8.1 presents eight different measures constructed from the resulting consumption expenditure data. The first four pertain specifically to these data. The first measure is the primary consumption expenditure measure, and the models used to analyse it only include our study's primary set of covariates correlated with our programme area dummy (p<.1). The second is this same measure but put on a log rhythmic scale (to normalize the distribution and reduce the influence of extreme observations) and analysed similarly. The next two measures are similar to the first two, but an index constructed from assets reportedly owned in 2007 weighted by the 2016 consumption expenditure data—described further below—are also controlled for. This was part of our larger effort to use asset recall as a means of predicting (estimating) baseline consumption expenditure levels.

Given the absence of baseline consumption expenditure data, we attempted to estimate what this would have been by weighting the assets owned by each household at this time by the 2016 consumption expenditure data.

The latter four measures presented in Table 4.8.1 are based on this approach, which deserves elaboration. As explained above, a household's wealth status in low and middle income countries is often measured using consumption expenditure data. An alternative way is by analysing household assets and other wealth indicators, e.g. the material of a household's roof and/or floor. Such data are commonly analysed using principal component analysis (PCA) or factor analysis. These are data reduction techniques that narrow in on the variation in household asset ownership, which is assumed to represent wealth status. The more an asset is correlated with this shared variation, the more weight it is given. Hence, each household's weighted index score is determined by both (a) the number of assets it owns; and (b) the particular weight assigned to each asset. This enables the *relative* wealth status between groups of households to be compared. However, such approaches suffer from a lack of theory to motivate either the choice of variables or the appropriates of derived weighting (O'Donnell et al. 2008).

[7] Adjusting for PPP was done to take into account Kenya's idiosyncratic purchasing power, i.e. the quantity of currency required to purchase a given basket of goods and services. The PPP conversion rates used were taken from the World Bank's website: http://data.worldbank.org/indicator/PA.NUS.PPPC.RF

[8] While dividing the above by household size as the overall denominator is recommended in the literature, it is considered important to avoid underestimating expenditure for larger sized households relative to their smaller counterparts. A recommended formula for computing household size for this purpose is: HH size $= (A + \alpha K)^{\theta}$ where *A* is number of adults in the household; *K* is the number of children; α is the cost of a child relative to an adult; and θ controls the extent of economies of scale. For low income countries, is recommend that α be set at .25 or .33 and θ be set at .9 (Deaton and Zaidi 1999).

TABLE 4.8.1: HH Consumption Expenditure and Asset Indices Weighted by these Data

	2016 HH consumption expenditure (CE)/capita				*Asset indices weighted by 2016 CE data*			
	16 CE w/ covariates only	16 CE log w covariates only	16 CE w/ CE 07 asset index	16 CE log w/ CE 07 asset index (log)	16 CE asset weighted w/ CE 07 asset index	16 CE asset weighted w/ CE 07 asset index (log)	Differenced 16 CE asset weighted	Differenced 16 CE asset weighted (log)
Raw results								
Sample Mean	4.58	1.37	4.58	1.37	4.92	1.48	0.95	0.20
Program Mean	4.61	1.39	4.61	1.39	5.01	1.49	1.00	0.21
Non-Program Mean	4.55	1.36	4.55	1.36	4.83	1.46	0.91	0.20
Unadjusted dif.	0.057	0.022	0.057	0.022	0.18*	0.031*	0.086	0.012
	(0.11)	(0.020)	(0.11)	(0.020)	(0.099)	(0.018)	(0.056)	(0.010)
Observations	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**
Raw results								
Sample Median	3.88	1.36	3.88	1.36	4.32	1.46	0.72	0.19
Program Median	3.91	1.36	3.91	1.36	4.35	1.47	0.74	0.20
Non-Program Median	3.85	1.35	3.85	1.35	4.27	1.45	0.69	0.18
Unadjusted dif.	0.059	0.015	0.059	0.015	0.082	0.019	0.047	0.023*
	(0.091)	(0.023)	(0.091)	(0.023)	(0.10)	(0.024)	(0.050)	(0.012)
Observations	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**
OLS (ITT)								
Coefficient	0.074	0.030	0.062	0.028*	0.10*	0.017*	0.10*	0.017*
	(0.10)	(0.019)	(0.087)	(0.017)	(0.055)	(0.0097)	(0.055)	(0.0098)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
IPWRA (ITT)								
Coefficient	0.087	0.029*	0.072	0.027*	0.084	0.015*	0.087*	0.013
	(0.085)	(0.016)	(0.084)	(0.016)	(0.054)	(0.0090)	(0.053)	(0.0092)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
Robust Reg. (ITT)								
Coefficient	0.14**	0.034*	0.13**	0.032*	0.065	0.019**	0.066	0.018*
	(0.072)	(0.019)	(0.065)	(0.017)	(0.043)	(0.0097)	(0.043)	(0.0098)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
Nearest Neighbor 1 to 1 Matching (ITT)								
Coefficient	-0.072	0.00037	0.038	0.020	0.12*	0.022	0.12*	0.023**
	(0.12)	(0.023)	(0.099)	(0.020)	(0.072)	(0.014)	(0.060)	(0.011)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
Quintile Reg. (ITT)								
Coefficient	0.12	0.028	0.16*	0.042**	0.077	0.019	0.099*	0.020*
	(0.092)	(0.023)	(0.088)	(0.021)	(0.058)	(0.012)	(0.054)	(0.012)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
2SLS (LATE)								
Coefficient	0.10	0.041	0.085	0.041	0.13	0.022	0.13*	0.022
	(0.16)	(0.030)	(0.13)	(0.026)	(0.15)	(0.025)	(0.079)	(0.014)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level for OLS and 2SLS models; covariates correlated with programme area (p=<0.1) used in all models; IPWRA models also include covariates correlated with outcome (p=<0.1) derived via stepwise regression; VSZ dummies used for fixed effects in all linear models, while exact matching within VSZ enforced for nearest neighbour matching models.

Hence, another approach to constructing such asset indices involves weighting each asset item based on the extent it is correlated with consumption expenditure, with the resulting coefficient estimates for each asset used as weights. Such consumption regression approaches attempt to predict consumption when complementary consumption data are available (O'Donnell et al. 2008). To implement this approach, we first took our binary set of 91 asset measures for 2016 (see Subsection Annex 2) and regressed them on the abovementioned 2016 daily household consumption measure. We then removed those assets negatively correlated and/or only modestly correlated with consumption expenditure (i.e. $p > 0.2$). This enabled us to arrive at a list of binary assets measures for the 2016 period that are all positively and at least moderately correlated with 2016 consumption expenditure. This set explains over 40% of the variation of the consumption data (R^2=0.4217). We then took this same set of assets owned during the 2007 period and gave each the same weights. Finally, we divided each index by the adjusted household size measure described in footnote 7. This study's primary outcome measure, which is declared in our pre-analysis plan,

takes the difference in these two consumption weighted indices between the baseline and endline periods.

The results presented in Table 4.8.1 reveal that Vi's programme does not seem to have significantly impacted these various consumption expenditure related measures, at least overall. However, the robust regression coefficients for consumption expenditure are statistically significant at the 5% and 10% levels, as are several of the consumption weighted asset measures. Estimates for the study's primary outcome variable—the differenced 2016 consumption weighted asset measure—range from 0.067 (robust regression) to 0.13 (2SLS). While all but the former are statistically significant at the 10% level, these effect sizes are very modest (*d*=0.0013, for example, for the OLS 0.10 estimate).

Our sub-group analyses for these measures (see Table A5.7, Annex 5) did, again, reveal significant variation. Several of the effect sizes for consumption expenditure for the Sirisia/Malakisi VSZ, for instance, are significantly larger than the overall averages. We applied robust regression to check the robustness of these results, and the estimated effects turned out to be even larger and more statistically significant. For example, for the single difference consumption expenditure model which controls for our standard covariates plus our 2007 asset consumption weighted asset measure, the estimated effect size from the robust regression model is 0.347 (*p*=0.007) against 0.132 (*p*=0.040) for the overall sample.

Overall, there is little evidence that Vi's programme significantly bolstered consumption expenditure. However, there is some evidence that it may have done so at the sub-group level.

However, this VSZ specific differential effect estimates in favour of the Sirisia/Malakisi do not transfer over to the consumption weighted asset measures. Here, it is the Tongaren/Likuyani households of the programme area that are better off and appear to have gained more over time for these particular measures. This is visually clear from an examination of the box plots presented in Figure 4.8.1. However, the estimated effect size for our primary outcome variable for this VSZ reduces considerably with the application of robust regression—0.31 (*p*=0.021) to 0.19 (*p*=0.025) but still remains statistically significant and much larger than those of the other VSZs.

However, the subgroup differential effect estimates that surprised us the most are for gender. In our pre-analysis plan, we hypothesized that female farmer group members would be less impacted by Vi's programme than their male counterparts. However, there is evidence that supports the opposite. This can be seen visually by the box plots presented in Figure 4.8.2 depicting our differenced consumption weighted asset index. While the OLS models presented in Table A5.7 reveal little difference between men and women, the visual difference between these box plots promoted us to investigate the potential for gender specific subgroup effects using robust and quantile regression, and the results are also presented in this same table. For the robust regression models, the effect sizes for women are again quite modest but all are significant below the 5% level, and their non-logged coefficients are significantly different from that of their male counterparts statistically speaking. For the quantile regression model, the differences are even starker, with the coefficients of all four models being statistically significant along gender lines. The median estimated gain for women, for example, on the study's primary outcome measure is 0.18 (*p*=0.009), against 0.099 (*p*=0.069) for the overall sample and -0.05 (*p*=0.556) for men.

Figure 4.8.1: Differenced Consumption Weighted Asset Index
By Program Area and Village Sampling Zone

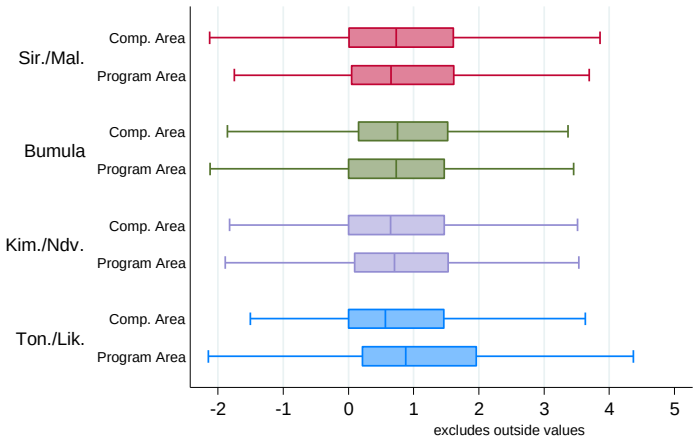
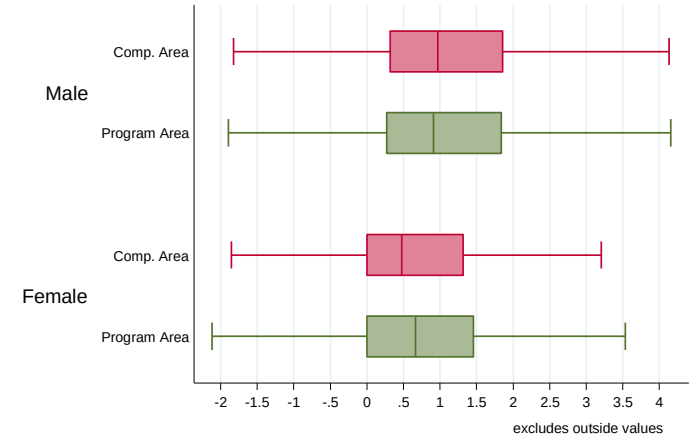


Figure 4.8.2: Differenced Consuption Weighted Asset Index
By Program Area and Gender



We found greater estimated programme effects for our primary outcome variable among households with female—as opposed to male—farmer group representatives.

Recall also that we hypothesized that dairy farmers would be impacted by Vi's programme more than non-dairy farmers. Yet, in the analyses thus far for agroforestry adoption and most of the non-dairy related intermediary outcomes, we have found the opposite to be the case. As revealed in Table A5.7, this is also true with respect to the consumption expenditure data but does not clearly cross over to the consumption weighted asset measures. Nevertheless, the consumption expenditure related differential effect estimates in favour of non-dairy farmers continue to persist following the application of both robust and quantile regression, as also show in Table A5.7.

One reason why the results for the consumption expenditure and consumption weighted asset measures are not necessarily similar is that they are measuring slightly different things. In general, consumption expenditure data are recognised as being more sensitive to recent changes in household income, while asset based measures typically reflect a household's more established wealth status (Moser and Felton 2007). This is because increases in household income generally need to be sustained for some time before they translate into significant increases in household asset ownership. In other words, after a household's income has been sustainably

increased, it will take time for it to accumulate assets, make improvements to the home, and so forth.

We therefore complemented our analysis of the 2016 pure consumption and consumption weighted asset measures with several other asset measures derived through PCA and what is dubbed the 'arbitrary' or 'naïve' approach (O'Donnell et al. 2008). The latter involves simply adding together the asset binary measures without differentially weighting them. We implemented the latter in part as a robustness check and in part to explore whether any differences between the programme and non-programme areas were particular to a specific asset class, e.g. housing characteristics or livestock ownership.

Our analyses of the consumption expenditure and consumption weighted asset measures was complemented with several PCA derived and 'raw' asset measures.

To implement PCA, we first took the binary asset for each time period and assessed their inter-item correlation and removed those negatively correlated with the other assets. The resulting inter-item correlation that resulted was quite high (alpha = 0.9025 and 0.8917 for the 2016 and 2007 binary asset measures, respectively).[9] We then constructed tetrachoric matrices with them, and principal component factor analysis was then run on these matrixes. Variables based on the first principal component were subsequently constructed. We did this for each time period for the overall dataset. We also repeated the above steps for each VSZ separately as a complementary measure that is specifically sensitive to the intra-VSZ variation in asset ownership.

Given that implementing PCA for each time period separately would generate different time-specific sets of asset weights, we avoided simply differencing the two indices to obtain a differenced measure. Rather, we first identified whether there had been gains over the two time periods for each asset indicator. Then we checked the inter-item correlation again for the resulting set, while iteratively removing negative values. We did this until we arrived with a low but still reasonable alpha of 0.7343 for the overall dataset. We then again constructed a tetrachoric matrix and ran principal component factor analysis on it again, thereby creating an 'asset gain index'. We further implemented the same steps above for each VSZ.

Table 4.8.2 presents our results for the overall dataset. It is clear that the programme effect estimates are more robust and consistent for the differenced PCA index and overall raw score asset measure. One likely explanation for this is that both narrow in on asset gains and hence are more sensitive to picking up such gains. However, these effect sizes are still modest, e.g. d=0.0037 for the differenced overall PCA measure. Nevertheless, the fact that the robust and quantile regression estimates are similar in size to those of the others gives us confidence that the results are not being driven by influential observations, i.e. they apply to the bulk of the distribution. Moreover, by examining the second half of the table, we see that the positive results in favour of the programme areas are not driven by any particular asset class; the relative gains seem to be across the board.

[9] When items are used in a scale or index, they should all measure the same underlying latent construct (e.g. household wealth status). The items, therefore, must be significantly correlated with one another. Cronbach's alpha is a measure of this inter-item correlation. The more the variables are correlated, the greater is the sum of the common variation they share. If all items are perfectly correlated, alpha would be 1 and 0 if they all were all independent from one another. For comparing groups, an alpha of 0.7 or 0. 8 is considered satisfactory (Bland and Altman 1997).

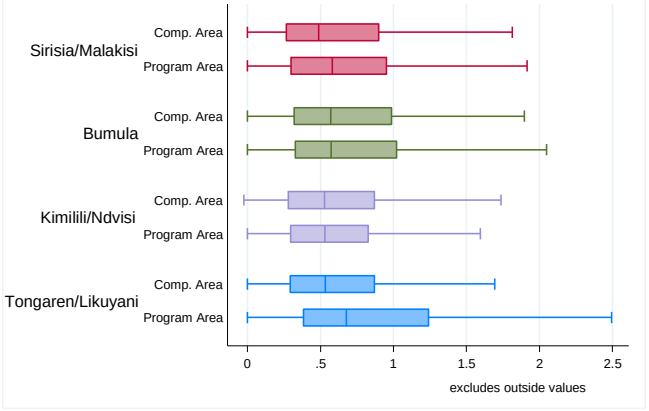
TABLE 4.8.2: Household Asset Wealth and Accumulation

	Principal Component Asset Indices				***Raw Asset Positive Gains***				
	2016 Overall	2016 by VSZ	Differenced Overall	Differenced by VSZ	Overall	House Char.	Home Durables	Productive Assets	Livestock
Raw results									
Sample Mean	2.01	2.12	0.72	0.67	8.52	1.25	3.88	1.72	1.68
Program Mean	2.05	2.16	0.75	0.69	8.78	1.32	3.95	1.76	1.75
Non-Program Mean	1.96	2.07	0.69	0.65	8.26	1.17	3.81	1.67	1.61
Unadjusted dif.	0.087*	0.085*	0.063***	0.038*	0.52***	0.15**	0.14	0.099	0.13**
	(0.046)	(0.048)	(0.022)	(0.021)	(0.19)	(0.067)	(0.095)	(0.061)	(0.059)
Observations	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**
Raw results									
Sample Median	1.76	1.87	0.56	0.53	8.00	1.00	4.00	1.00	1.00
Program Median	1.80	1.92	0.59	0.54	8.00	1.00	4.00	1.00	1.00
Non-Program Median	1.70	1.81	0.53	0.51	7.00	1.00	4.00	1.00	1.00
Unadjusted dif.	0.10**	0.10*	0.061***	0.029	1***	0	0	0	0
	(0.052)	(0.055)	(0.023)	(0.020)	(0.21)	(0.14)	(0.14)	(0.14)	(0.14)
Observations	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**
OLS (ITT)									
Coefficient	0.050*	0.040	0.069***	0.048**	0.59***	0.16**	0.17*	0.14**	0.12*
	(0.026)	(0.028)	(0.025)	(0.023)	(0.21)	(0.071)	(0.10)	(0.070)	(0.068)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
IPWRA (ITT)									
Coefficient	0.045**	0.033	0.061***	0.041**	0.54***	0.15**	0.14	0.14**	0.12**
	(0.022)	(0.023)	(0.021)	(0.020)	(0.18)	(0.067)	(0.088)	(0.060)	(0.059)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
Robust Reg. (ITT)									
Coefficient	0.053**	0.044**	0.050***	0.047***	0.57***	0.063	0.20**	0.14**	0.11*
	(0.021)	(0.022)	(0.017)	(0.016)	(0.17)	(0.042)	(0.088)	(0.055)	(0.056)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
Nearest Neighbour 1 to 1 Matching (ITT)									
Coefficient	0.051	0.045	0.050**	0.045**	0.48**	0.071	0.087	0.19***	0.13*
	(0.032)	(0.034)	(0.024)	(0.023)	(0.21)	(0.075)	(0.11)	(0.070)	(0.074)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
Quintile Reg. (ITT)									
Coefficient	0.062***	0.054**	0.056***	0.035*	0.73***	0.069	0.26**	-3.0e-16	0.092
	(0.023)	(0.026)	(0.020)	(0.019)	(0.20)	(0.11)	(0.10)	(0.056)	(0.094)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
2SLS (LATE)									
Coefficient	0.064*	0.051	0.087***	0.061**	0.75***	0.20**	0.22*	0.18**	0.15*
	(0.033)	(0.035)	(0.032)	(0.029)	(0.27)	(0.091)	(0.13)	(0.088)	(0.086)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level for OLS and 2SLS models; covariates correlated with programme area (p=<0.1) used in all models; IPWRA models also include covariates correlated with outcome (p=<0.1) derived via stepwise regression; VSZ dummies used for fixed effects in all linear models, while exact matching within VSZ enforced for nearest neighbour matching models.

Table A5.8 in Annex 5 presents the results of our subgroup analysis for these particular asset measures. There is, yet again, variation across the VSZs, with the effect estimates for Tongaren/Likuyani standing out as being much greater than the others. Figure 4.8.3 presents box plots for each VSZ for the differenced asset PCA measure, which clearly reveals this. However, the distributions for Sirisia/Malakisi also indicate differences, albeit less pronounced, despite the OLS results indicating otherwise. Indeed, robust and quantile regression generate effect estimates for this VSZ of 0.066 (p=0.046) and 0.091 (p=0.021), respectively.

Figure 4.8.3: Asset Gain Index (PCA 1st Component)
By Program Area and Village Sampling Zone



The differential effects in favour of households with female participants cross over to these measures as well, while this is not the case for non-dairy producers. Table A5.8 shows this quite clearly, with all the effect estimates for women being highly significant but with virtually no evidence of any impact for men. However, our Wald tests did not find the effect sizes to be significantly different from that of men in a statistical sense for the variations of our asset gain measure. However, this is the case for the two single difference PCA measures and the overall raw asset measure. Moreover, our implementation of robust and quantile regression models confirm the robustness of the results. Figure 4.8.4 provides a nice visual to illustrate the differences in the distributions for the differenced asset measure. Male participant households are better off overall, but those represented by women in the programme area gained more than their counterparts in the non-programme area.

The positive differential effect estimates identified for households with female farmer group members for the consumption weighted asset measures transfers over to the other asset based measures.

Figure 4.8.4: Asset Gain Index (PCA 1st Component)
By Program Area and Gender

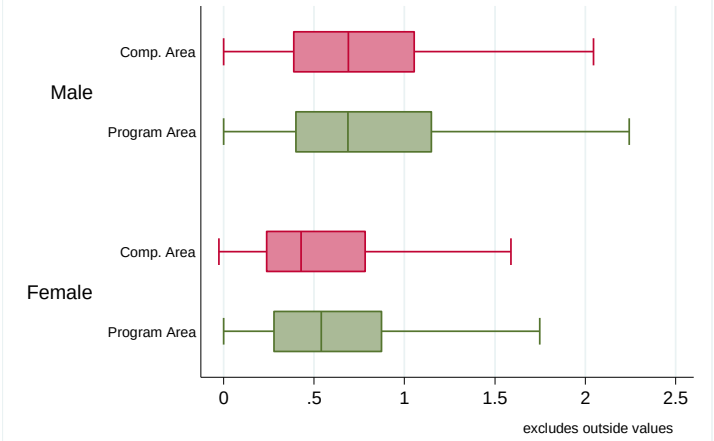
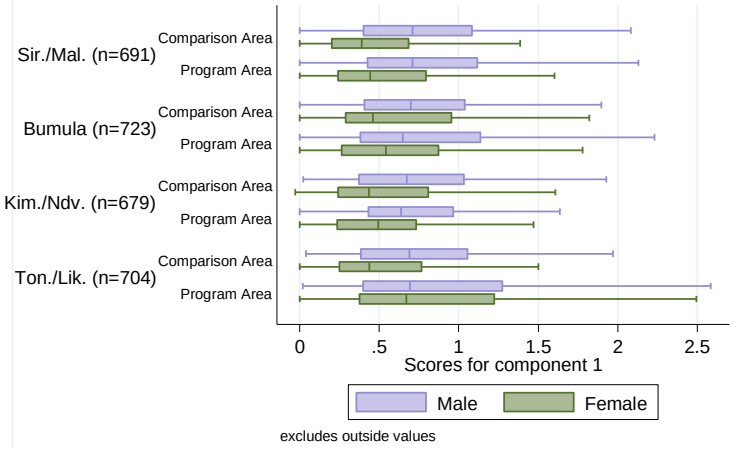


Figure 4.8.5 complements Figure 4.8.4 by showing a breakdown of the distributions for this same measure by VSZ and gender. While the differences is certainly greatest in the Tongaren/Likuyani VSZ, there is a common pattern across all VSZs, particularly for the median values.

Figure 4.8.5: Asset Gain Index (PCA 1st Component)
By Program Area, Gender, and Village Sampling Zone



The differential asset-based effect estimates in favour of households with female farmer group members stood up to several additional robustness tests.

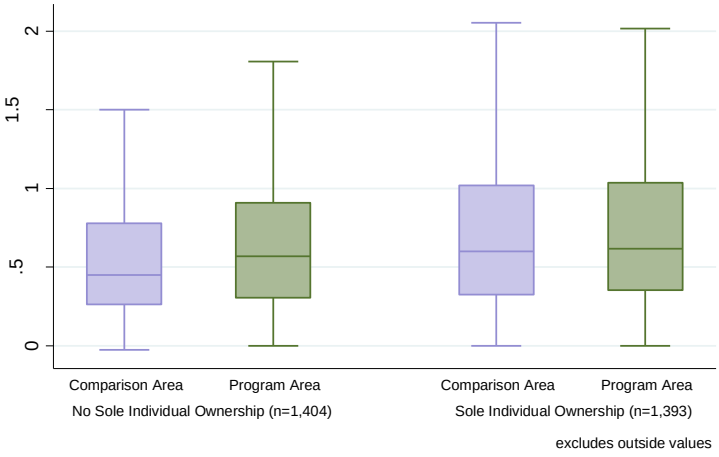
The covariates we used in our various subgroup analyses are the same ones used in our models for the overall dataset, i.e. those correlated with being in the programme area (p<0.1). However, this overall set of covariates may not necessarily reflect subgroup specific covariate correlation with our programme area dummy variable. Consequently, as a further robustness check of the results for households with female farmer group members we, therefore, replicated several analyses again with a set of covariates correlated with being a female farmer group member in the programme area (p<0.1). (See Annex 6 for these covariates.) For the consumption weighted asset measure, the effect estimate (0.095) was, like that of the original model, statistical insignificant (p=0.175). However, the robust and quantile regression estimates are at 0.13 (p=0.018) and 0.20 (p=0.001), respectively. For the asset gain measure, the results are consistent with those of the original model as well: 0.09 (p=0.001) for the OLS model, 0.06 (p=0.004) for the robust regression model, and 0.08 (p=0.000) for the quantile regression model.

While not in our registered pre-analysis plan, we also generated another surprising and noteworthy subgroup effect estimate. At the beginning of the land parcel survey module, the respondent was asked about the current and past ownership arrangements associated with this parcel, with options such as self with title deed; self with no title deed; self and spouse together with title deed; etc. From the responses, we created several variables pertaining to the baseline period, including (A) land owned by self *without* spouse and with or without formal land title (50% of the sample); (B) land owned by self *with or without* spouse and with or without formal land title (58% of the sample); and (C) household with formal land title (33% of the sample). The results of the sub-group analysis for sub-group A are presented in Table A5.8. Oddly, households with respondents who stated they did not solely own the land from the programme area gained more on the asset gains index over time vis-à-vis their counterparts in the comparison area. For the sub-group that reported owning their household's parcel outright (sub-group A), there appears to be no difference in the programme and comparison areas in terms of asset accumulation from the baseline period. Figure 4.8.6 reveals this visually. The results are similar for subgroup B, but are not as stark, and, for subgroup C, there are no significant differential effects for these asset measures. Note that the specific fixed effects of the gender of the respondent and VSZ were controlled for in all models, and these same differential

effects for subgroups A and B do not exist for the consumption and consumption weighted asset measures.

Oddly, households with farmer group members who do not own their household's main parcel are estimated to have experienced greater gains in asset accumulation from Vi's programme.

Figure 4.8.6: Asset Gain Index (PCA 1st Component)
By Programme Area and Individual Parcel Ownership



In this subsection we explored the extent to which there is evidence that Vi's programme positively impacted overall household wealth, as measured by consumption expenditure and several indices based on assets and other household wealth indicators. There is little evidence that Vi's programme significantly bolstered household consumption expenditure—at least overall. However, this may have taken place in the Sirisia/Malakisi VSZ and among non-dairy farmers. Given the absence of baseline data on consumption expenditure, for this to be true we would need to assume that (a) their respective comparators were at similar consumption levels at baseline and/or our consumption weighted asset measure for 2007 and other covariates effectively controlled for any relevant baseline related differences that may have existed; and (b) both subgroups in the programme and comparison villages were subjected to similar non-programme related events and trends.

While modest, the evidence supporting the overall impact of Vi's programme on asset accumulation is more robust and convincing. This is particularly the case for our differenced PCA asset gain measure, rather than our differenced consumption weighted asset measure that we declared as this study's primary outcome variable. However, even with respect to our asset gain measure, we found considerable variation among particular subgroups. The programme villages in the Tongaren/ Likuyani VSZ and, to a lesser extent, the Sirisia/Malakisi VSZ, for example, experienced greater asset accumulation vis-à-vis their comparators in the comparison villages, a trend that does not seem to have taken place in the other VSZs. Moreover and surprisingly, households with female farmer group members and those with members who do not themselves own their respective household's main agricultural parcel in the programme area experienced greater asset accumulation over time, while households with male members and members that own (irrespective of formal title) their household's main parcel outright in this same programme area did not, at least, on average. We will now examine several other complementary outcome measures to assess the extent to which the results are consistent with those presented above.

4.9 Coping Strategy Index and Food and Nutritional Security

In addition to improving overall household wealth status, our theory of change for VI's programme postulates that the uptake of the promoted agroforestry tree species and practices bolstered both resilience to shocks and household food and nutritional security. To measure the former, the respondents were first asked whether their local area had experienced a serious drought, irregular rains, flooding, landslides, and/or a crop or livestock related pest or disease outbreak over the past three years. If they responded in the affirmative to any of these categories, they were then asked the extent to which their households were affected. If they reported at least to some extent, they were coded as being part of the sub-sample of households that had experienced one or more shocks in the last three years. For each shock type, the respondent was then asked what strategies their household used to deal with these shocks based on the Coping Strategies Index (Maxwell and Caldwell 2008). These strategies are presented in Figure 4.9.1, including the weights assigned for each strategy. Skipping an entire day without eating, for example, is assumed to be a more extreme coping mechanism than substituting commonly bought foods with those of a cheaper kind and, hence, is given relatively more weight.

We assume that households more resilient to shocks will need to resort to less drastic coping strategies when they take place and, hence, will have lower scores on the Coping Strategies Index.

Figure 4.9.1 also presents the percentages of respondents who reported that their respective households had resorted to each particular coping strategy. We see that reducing the quantity and/or types of food consumed are the most popular strategies, followed by borrowing money or food. The only clear significant differences between the programme area and Vi group households on the one hand and the comparison area households on the other is with respect to the mortgaging or the selling off of assets, with households in the former area reporting this to a significantly lesser extent. This prompted us to analyse this coping response independently from the overall index.

Figure 4.9.1: Reported Coping Strategies for HHs at Least Partially Impacted by Shock Over Previous 3 Years (n=2014)

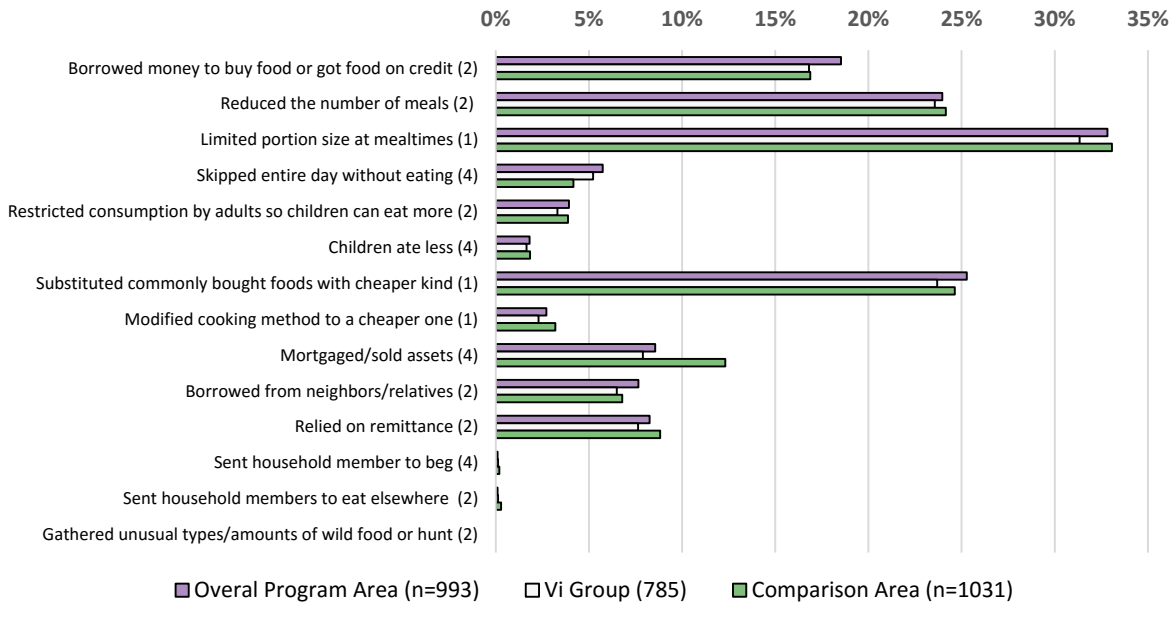


TABLE 4.9.1: Coping Strategy Index & Food Security Measures

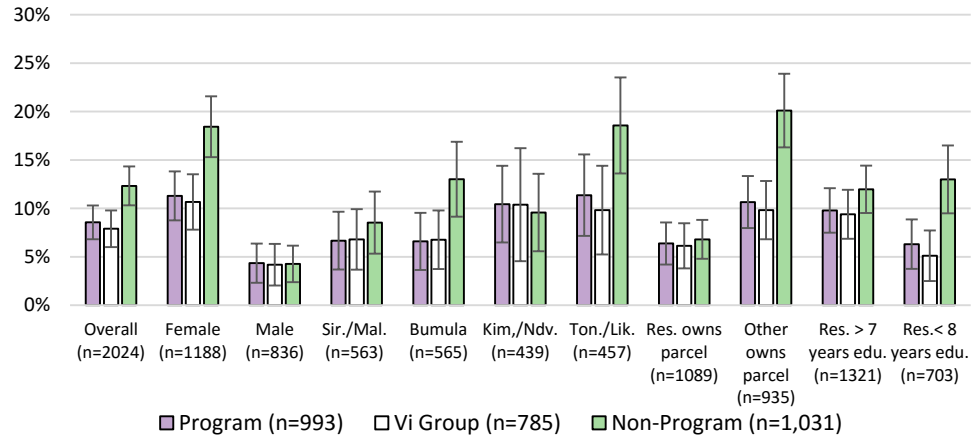
	Coping Strategy Index			***MDD-W***			***Food Shortage Months***		
	Full Score	> 2 points	Distressed Asset Sale	Score/10	>4 points	Meat	Last 12 months	Dif. from baseline	Reduction from baseline
Raw results									
Sample Mean	2.54	0.40	0.10	5.20	0.59	0.44	2.21	-0.21	0.33
Program Mean	2.51	0.39	0.09	5.27	0.61	0.45	2.16	-0.24	0.32
Non-Program Mean	2.57	0.41	0.12	5.13	0.57	0.43	2.25	-0.18	0.33
Unadjusted dif.	-0.061	-0.022	-0.038***	0.15**	0.038**	0.024	-0.089	-0.055	-0.0058
	(0.13)	(0.022)	(0.014)	(0.068)	(0.019)	(0.019)	(0.073)	(0.068)	(0.018)
Observations	**2024**	**2024**	**2024**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**
Raw results									
Sample Median	2.00	0.00	0.00	5.00	1.00	0.00	2.00	0.00	0.00
Program Median	2.00	0.00	0.00	5.00	1.00	0.00	2.00	0.00	0.00
Non-Program Median	2.00	0.00	0.00	5.00	1.00	0.00	2.00	0.00	0.00
Unadjusted dif.	--	--	--	--	--	--	--	--	--
Observations	**2024**	**2024**	**2024**	**2797**	**2797**	**2797**	**2797**	**2797**	**2797**
OLS (ITT)									
Coefficient	-0.090	-0.073	-0.25**	0.15	0.11*	0.060	-0.088	-0.055	-0.014
	(0.16)	(0.068)	(0.12)	(0.096)	(0.063)	(0.059)	(0.089)	(0.074)	(0.055)
Observations	**2018**	**2018**	**2018**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
IPWRA (ITT)									
Coefficient	-0.11	-0.032	-0.040***	0.15**	0.042**	0.023	-0.082	-0.049	-0.0039
	(0.12)	(0.021)	(0.012)	(0.066)	(0.018)	(0.019)	(0.069)	(0.069)	(0.018)
Observations	**2018**	**2018**	**2018**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
Robust Reg. (ITT)									
Coefficient	-0.05	--	--	0.14*	--	--	-0.13*	-0.0023	--
	(0.12)			(0.071)			(0.072)	(0.054)	
Observations	**2018**			**2790**			**2790**	**2790**	
Nearest Neighbour 1 to 1 Matching (ITT)									
Coefficient	-0.25*	-0.055**	-0.044***	0.20**	0.035	0.027	-0.11	-0.0062	-0.019
	(0.15)	(0.025)	(0.015)	(0.082)	(0.022)	(0.023)	(0.086)	(0.085)	(0.022)
Observations	**2018**	**2018**	**2018**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
Quintile Reg. (ITT)									
Coefficient	-0.12	--	--	0.053	--	--	-0.11	No	--
	(0.17)			(0.088)			(0.10)	cover.	
Observations	**2018**			**2790**			**2790**		
2SLS (LATE)									
Coefficient	-0.12	-0.098	-0.32**	0.19	0.14*	-0.11	-0.069	-0.018	-0.12
	(0.20)	(0.085)	(0.14)	(0.12)	(0.080)	(0.11)	(0.094)	(0.070)	(0.20)
Observations	**2018**	**2018**	**2018**	**2790**	**2790**	**2790**	**2790**	**2790**	**2018**

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level for OLS and 2SLS models; covariates correlated with programme area (p=<0.1) used in all models; IPWRA models also include covariates correlated with outcome (p=<0.1) derived via stepwise regression; VSZ dummies used for fixed effects in all linear models, while exact matching within VSZ enforced for nearest neighbour matching models.

The first section of Table 4.9.1 presents the results of the comparison of the programme and comparison households with respect to the full Coping Strategy Index, a binary measure with a cut-off of greater than two points on this index, and the distressed asset sale binary measure. Both the raw and model based estimates yielded no significant differences for the overall index and its binary variation. However, about 4% to 5% fewer households in the non-programme area reported distressed asset sale as a coping strategy. Given that 2,024 out of our sample of 2,797 households reported being affected by a shock during the previous three years, the covariates used in the various models, similar to those used for dairy farmers above, are based on this sub-sample's correlation with being in the programme area (p<0.1).

Our subgroup analysis for the coping strategies measures (Table A5.9 in Annex 5) revealed no differential programme effects for the first two measures. However, there are several noteworthy ones for the distressed asset sale binary measure, as indicated in Figure 4.9.2. While female respondents were more likely to report such distressed sales than males in general, this is significantly greater in the comparison area. This trend is also similar for the Tongaren/Likuyani and Bumula VSZs and among those respondents that do not own their household's main farming parcel and have less education.

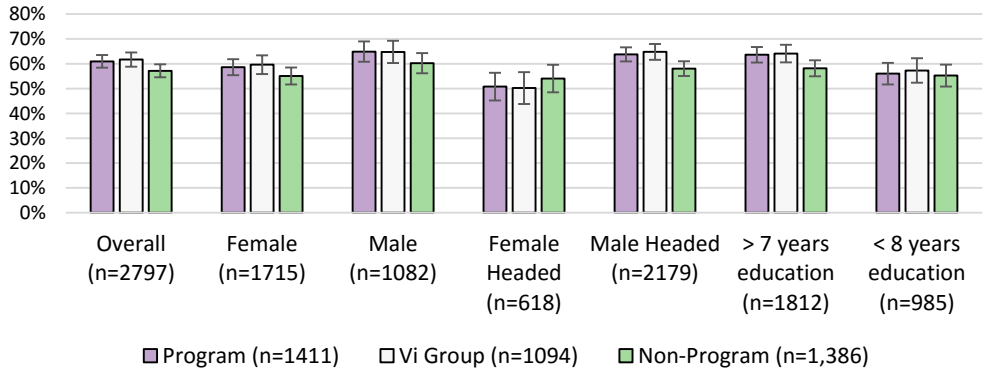
Figure 4.9.2: % of respondents reporting distressed asset sale by gender, VSZ, parcel ownership & education



There are no significant overall or sub-group specific differences for the overall Coping Strategies Index. However, the case is different for its embedded distressed asset sale indicator.

For household food security, our survey captured data on two indicators: the Minimum Dietary Diversity for Women (MDD-W) (FAO and FHI 360 2016) and Months of Adequate Household Food Provisioning (Bilinsky and Swindale 2010). For the former, the respondents were asked if they had consumed various food items during the previous day from a list of 17 such items, which we later grouped into MDD-W's 10 food group categories. We analysed the resulting data as a continuous score out of 10 and as a binary variable using this measure's 'official' cut-off of five or more points. We also assume the consumption of meat is the context of this study could indicate higher wealth status, so we analysed this variable separately as well. As indicated in Table 4.9 only several of the models yielded statistically significant results in favour of the programme area and, again, these effect sizes are very modest.

Figure 4.9.3: Minimum Dietary Diversity Women (≥ 5/10 food groups)

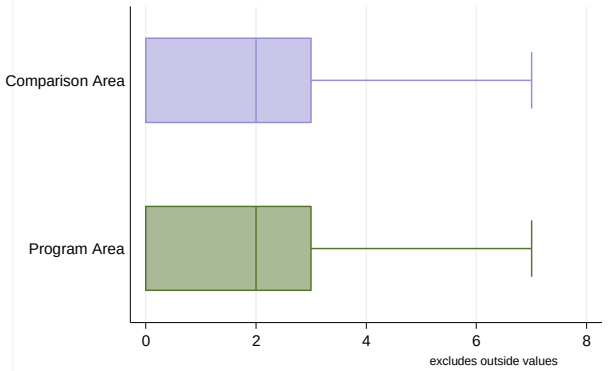


We found little evidence that Vi's programme significantly impacted our two measures of household food and nutritional security, both overall and at the sub-group level.

Our sub-group analyses of the MDD-W data (Table A5.9) only yielded several modest differential sub-group effect estimates as well. While there were no statistically significant differential effects identified among male and female respondents, a modest one was identified in favour of male headed households. Also, a slightly higher percentage respondents with greater years of education reached the MDD-W's binary threshold in the programme area vis-à-vis their counterparts in the comparison area.

To capture data on the second food security measure—months of adequate food provisioning—the enumerators asked the respondents the following question: "Were there months in the past 12 months—that is since August of last year—in which you did not have enough food to meet your family's needs?" If they responded in the affirmative, they were asked to recall the particular months. They were also asked to recall this for the baseline year. As indicated in Table 4.9.1 and by the box plots in Figure 4.9.4, no significant differences between the programme and comparison sites were found. Our sub-group analyses (Table A5.9) also did not reveal any significant programme effect estimates at the subgroup level.

Figure 4.9.4: # of Months with Reported Food Shortages in Last 12 Months By Programme Aea



4.10 Education Progression, Spending & the Economic Ladder

During our formative qualitative work, which included several focus group discussions with Vi group members, we were told that having trees on farm facilitates the making of lump sum expenditures, e.g. paying for school and hospital fees, through the sale of timber, fuelwood, and other agroforestry products. This motivated us to construct and analyse several variables. One relates to education progression: If Vi group households had established more trees on their farms, it is possible that they were in a better position to send their children to school in general and secondary school in particular. Table 4.10 presents three different variables pertaining to this: % of 14-17 year old children who progressed on to secondary school and % of households where any and all 14-17 year old children have done so, respectively. As is clear, the households in the *comparison area* are actually slightly better off in relation to these variables, particularly the second binary variable.

Table A5.10 presents the results of our sub-group analysis for these education progression measures, while Figure 4.10 highlights of the percentages of households where all their 14 to 17 year old children have all progressed to secondary education in particular. As is clear, there is variation in the results. Approximately 4% fewer households from those represented by female famer group members scored positively on this measure vis-à-vis their counterparts in the non-programme area. However, this difference is statistically insignificant. Among the VSZs, the variation is even greater, particularly for Bumula where there is a large and statistically significant difference (14.6%) in favour of the households in the comparison area. In the Kimilili/Ndivisi VSZ, however, there is a 7.5% difference in favour of the programme area, which is

statistically significant at the 10% level.

TABLE 4.10: Education Progression and Spending & the Economic Ladder

	Secondary Progression for 14-17 children			***Education Expenditure per >5 Child***			***Economic Ladder***	
	% of children gone to sec.	Any child gone to sec.	All children gone to sec.	Ksh spent in previous year	Log of Ksh spent in previous year	Expenditure about median	Overall change	At least 1 step up
Raw results								
Sample Mean	0.71	0.79	0.62	17133.73	8.39	0.48	0.73	0.66
Program Mean	0.69	0.76	0.61	17107.67	8.42	0.48	0.75	0.66
Non-Program Mean	0.73	0.81	0.64	17160.15	8.35	0.48	0.72	0.66
Unadjusted dif.	-0.046**	-0.054***	-0.037	-52.5	0.075	-0.0041	0.031	0.0025
	(0.020)	(0.020)	(0.024)	(1148.5)	(0.091)	(0.020)	(0.038)	(0.018)
Observations	**1698**	**1698**	**1698**	**2533**	**2533**	**2533**	**2797**	**2797**
OLS (ITT)								
Coefficient	-0.032	-0.16**	-0.070	-650.0	0.063	-0.040	0.039	0.027
	(0.022)	(0.077)	(0.070)	(1349.6)	(0.11)	(0.059)	(0.044)	(0.057)
Observations	**1693**	**1693**	**1693**	**2526**	**2526**	**2526**	**2790**	**2790**
IPWRA (ITT)								
Coefficient	-0.027	-0.044**	-0.015	No conv.	0.12	-0.0075	0.036	0.0088
	(0.019)	(0.020)	(0.022)		(0.079)	(0.017)	(0.038)	(0.018)
Observations	**1693**	**1693**	**1693**	**2526**	**2526**	**2526**	**2790**	**2790**
Robust Reg. (ITT)								
Coefficient	No conver.	--	--	115.5	-0.0056	--	0.048	--
				(367.7)	(0.071)		(0.036)	
Observations				**2526**	**2526**		**2790**	
Nearest Neighbor 1 to 1 Matching (ITT)								
Coefficient	-0.030	-0.033	-0.026	-1132.7	0.034	-0.0076	0.077*	0.026
	(0.025)	(0.025)	(0.029)	(1352.7)	(0.11)	(0.024)	(0.045)	(0.022)
Observations	**1693**	**1693**	**1693**	**2526**	**2526**	**2526**	**2790**	**2790**
Quintile Reg. (ITT)								
Coefficient	No conver.	--	--	209.3	0.057	--	0	--
				(639.9)	(0.091)		(0.017)	
Observations				**2526**	**2526**		**2790**	
2SLS (LATE)								
Coefficient	-0.040	-0.21**	-0.086	-818.0	0.079	-0.050	0.049	0.034
	(0.027)	(0.096)	(0.088)	(1690.4)	(0.13)	(0.074)	(0.055)	(0.073)
Observations	**1693**	**1693**	**1693**	**2526**	**2526**	**2526**	**2790**	**2790**

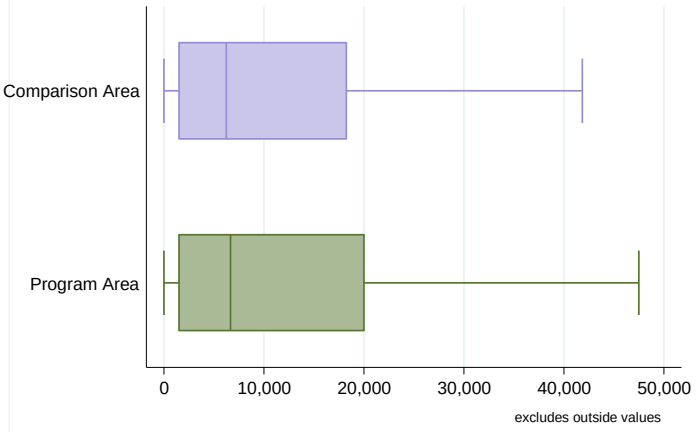
* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level for OLS and 2SLS models; covariates correlated with programme area (p=<0.1) used in all models; IPWRA models also include covariates correlated with outcome (p=<0.1) derived via stepwise regression; VSZ dummies used for fixed effects in all linear models, while exact matching within VSZ enforced for nearest neighbour matching models.

Households in the programme area were found to be slightly worse off for our educational progression measures, but with significant variation at the subgroup level.

Figure 4.10.1: Percentages of HHs with 14 to 17 Year Old Children who have All Progressed to Secondary Education by Programme Area, Gender, VSZ & Baseline Wealth Status

The second set of variables in Table 4.10 are in relation to educational spending. In particular, our theory of change expects households with more trees on farm are in a better position to either send more children to school, to better schools, etc., at least on average. As is clear from Table 4.10 and the box plots presented in Figure 4.10.2, we did not find any statistically significant differences between households in the programme and comparison areas with respect to such spending. Our subgroup analysis revealed some variation in this trend in favour of particular VSZs and non-dairy producers, but the results are statistically insignificant for the log transformed measure, revealing that the differences are likely driven by several influential observations.

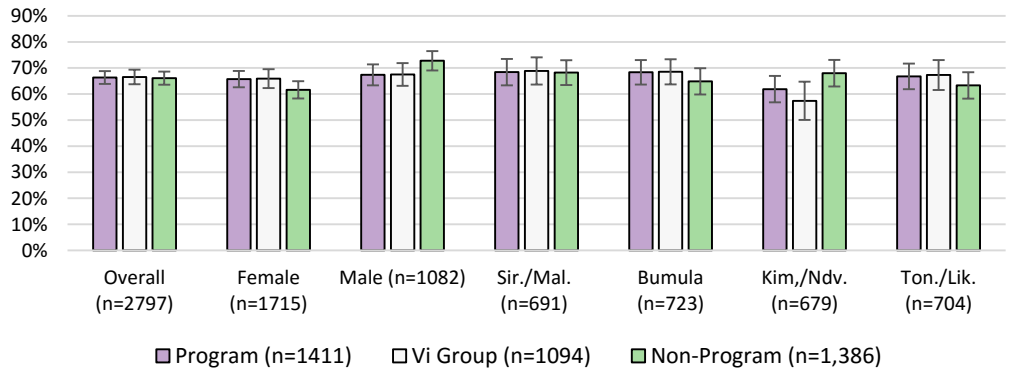
Figure 4.10.2: Amount Spent (Ksh) Per Household on Average Per Child By Programme Area



Overall, no significant differences between the programme and comparison areas were found for our Economic Ladder derived measures, but with some differences at the subgroup level.

The final set of measures presented in Table 4.10 pertain to an adapted version of the Economic Ladder Question (ELQ) (Ravallion and Lokshin 1999). To obtain data on these measures, the respondents were shown a drawing with six steps, and explained that the poorest farmers in the sub-county are on the bottom step, while the richest are on the top step. They were then asked to indicate the step their households have been over the past year, as well as during the baseline period. For the overall sample, 23% stated that they are on the same step now as in 2007, while 66% and 11% said that they had gone up and down one or more of the steps, respectively. As presented in Table 4.10, we analysed the differences in reported step placement between the time periods, as well as a binary measure indicating a reported movement of at least one step up on the ladder. As is clear, no significant differences were found between the programme and comparison areas.

Figirue 4.10.3: % of Respondents Reporting that their Households Have Progressed at Least One Step up the Economic Ladder



Our sub-group analysis (Table A5.10), however, did find several small, yet statistically, significant differences along gender and VSZ lines. As shown in Figure 4.10.3, slightly more women and less men in the programme area reported progressing at least one step up on the ladder, vis-à-vis their counterparts in the comparison area. There is further some variation among the VSZs, with the relative differences between the programme and comparison area between Bumula and Kimulili/Ndivisi, in particular, being statistically significant.

4.11 Investigating mechanisms for key greater asset accumulation among households in the Programme Area

As reviewed above, there is evidence that participation in Vi's programme resulted in greater (albeit modest) asset accumulation, particularly among households with female participants and in specific VSZs. It also seems to have reduced distressed asset sales as well, particularly among households with female participants. The purpose of this sub-section is to explore how Vi's programme may have worked to bring about these particular downstream effects. Was it because, as we have seen, households in the programme area adopted agroforestry and/or earned more income from the sale of agroforestry products to a greater extent than those in the comparison area? Or could it have been through some other dimension(s) of this programme, e.g. the promotion of other sustainable agricultural land management (SALM) practices and/or the promotion of group savings and lending activities?

We used regression and mediation analysis to interrogate four hypotheses on how the estimated programme effects on asset accumulation may have come about.

In what follows, we present hypothesized mechanisms for how the programme's estimated effects on asset accumulation and reduced distressed asset sales may have come about and explore—quantitatively—the extent each is consistent with the data. We then conclude by reviewing relevant findings from our qualitative data.

Hypothesis 1: The differential uptake of agroforestry in the programme area—as opposed to other dimensions of Vi's programme—was significantly responsible for the estimated effects on asset accumulation

Consistent with our theory of change for Vi's programme, we have seen that relatively greater (albeit modest) gains in both the AF index and plot-level tree cover took place in the overall programme area vis-à-vis the comparison area. To what extent could this modest uptake of agroforestry be responsible for the corresponding modest programme effect estimates for asset accumulation?

Table 4.11.1 presents the results of the analyses we undertook to explore this question. For each of our four central asset measures presented in column 1, the overall OLS derived effect estimates are presented in column 2. The coefficients presented in the third and fourth columns reveal how those of column 2 change with the inclusion of two possible mediating mechanisms, respectively: (a) agroforestry uptake, as measured with both the differenced AF index and the differenced plot-level tree cover measure; and (b) the uptake of the other two key components of Vi's programme, as measured by our differenced SALM practice index and a micro-enterprise participation index.[10] If the overall OLS coefficients change significantly

[10] We constructed this index in a similar way to both the AF index and other SALM practice index. It comprises of seven indicators weighted equally under the following three dimensions: (1) HH has off-farm business as livelihood; (2) Respondent is active in at least one micro-finance group, i.e. respondent is member, participates in decision-making to at least a medium extent, and attended a meeting at least 4 times in the previous 12 months; and (3) Respondent has been trained in a micro-finance/business topic and reports having had

with the inclusion of these potential intermediary mechanisms, there is support—but by no means conclusive proof—that mediation could have taken place in this way. If they do not change, mediation through the hypothesized mechanism is unlikely (assuming adequate construct measurement), given that mediation requires, at a minimum, for the mediator variable(s) to be correlated with both treatment and outcome (MacKinnon 2008). This is not the case if the overall OLS coefficient remains unchanged.

A review of the OLS coefficients presented in Table 4.11.1 reveal that the data supports the hypothesis that the adoption of agroforestry was responsible for the programme's estimated effects on asset accumulation to a greater degree than the other SALM practices and microfinance/enterprise components of Vi's programme. However, the coefficients for the distressed asset sale outcome variable do not change significantly for either of the two sets of mechanisms, save for households with female participants.

Table 4.11.1: Results of Mediation Analysis—Asset Effect Estimates through AF Index & Tree Cover

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		OLS Models		**Mediation Analysis Results for AF Uptake**					
	Overall coef.	**w/ AF & Tree indices**	**w/ SALM & Micro-enter.**	**Total Effect**	**Direct Effect**	**Indirect Effect**	**p̂ mediated total**	**p̂ mediated AF index**	**p̂ mediated trees cov.**
Overall									
Differenced Consump. Weighted Index	0.10*	0.049	0.094	0.111*	0.060	0.052***	0.465	0.380	0.085
	(0.062)	(0.062)	(0.061)	(0.059)	(0.059)	(0.012)			
PCA 2016 Asset Index	0.050*	0.028	0.046*	0.051**	0.029	0.023***	0.439	0.324	0.115
	(0.026)	(0.027)	(0.026)	(0.025)	(0.026)	(0.005)			
Differenced Asset Gain Index	0.069***	0.048*	0.065***	0.068***	0.047*	0.021***	0.304	0.196	0.107
	(0.025)	(0.026)	(0.025)	(0.025)	(0.025)	(0.01)			
Distressed Asset Sale	-0.033**	-0.032**	-0.030**	-0.042	-.0040	-0.002	0.056	0.054	0.002
	(0.015)	(0.015)	(0.014)	(0.019)	(0.019)	(0.003)			
HHs with Female Participants Only									
Differenced Consump. Weighted Index	0.095	0.044	0.083	0.089	0.038	0.051***	0.570	0.470	0.100
	(0.07)	(0.069)	(0.070)	(0.07)	(0.07)	(0.016)			
PCA 2016 Asset Index	0.086***	0.062**	0.079***	0.091***	0.067**	0.024***	0.266	0.207	0.059
	(0.029)	(0.030)	(0.029)	(0.029)	(0.029)	(0.01)			
Differenced Asset Gain Index	0.094***	0.074***	0.090***	0.095***	0.074***	0.021***	0.216	0.149	0.067
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.01)			
Distressed Asset Sale	-0.067**	-0.049**	-0.059**	-0.084***	-0.0686**	-0.015***	0.181	0.175	0.006
	(0.026)	(0.025)	(0.025)	(0.03)	(0.03)	(0.004)			

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with Programme Area used in OLS models; mediation analysis models also include covariates correlated with mediator variable measures. Mediation models are all saturated, i.e. they were constructed to perfectly reproduce all variances, co-variances, and means.

Given that these results favour agroforestry uptake as being a key intermediary mechanism, we interrogated this more formally via mediation analysis using Stata's *sem* command. The particular models we constructed include both covariates correlated with programme area and the assumed mediator variables. Column 6 of

implemented this training to at least a medium extent. These indices were constructed for both the 2016 and 2007 periods and differenced.

The results of our mediation analyses support the hypothesis that the uptake of agroforestry was significantly responsible for greater asset accumulation in the programme area.

Table 4.11.1—Total Effect—corresponds to the overall programme effect estimated by the mediation models, while the Direct Effect estimates (column 6) and Indirect Effect estimates (column 7) correspond to the extent to which (a) the independent variation of programme area dummy and (b) the variation it shares with the hypothesised mediating variable accounts for variation in the outcome measure, respectively. The proportion of the total effect mediated (column 8), then, corresponds to the proportion of the total effect that can be accounted for by the variation our programme area dummy shares with the hypothesised mediator variable. A significant indirect effect estimate and, in turn, a high proportion of total effect mediated does not, however, conclusively evidence mediation; it is possible that one or more unmeasured factors correlated with both may be responsible.

The results for Table 4.11.1 reveal that a significant share of our programme effect estimates can be explained by the uptake of agroforestry, i.e. the data are consistent with this hypothesised mechanism. However, there is still significant variation—over 50% for most models—that is not explained by our agroforestry uptake measures. It is also worth noting that, while the overall and female specific indirect effect estimates are similar to the overall indirect effect estimates for the first three measures, their respective proportions of total effect mediated differ because their overall direct effect estimates differ. Observe further that the indirect effect estimate for the distressed asset sale outcome measure is only significant for the female respondents. Finally, note that a significantly greater proportion of the programme's effect is mediated through the multidimensional differenced AF index measure, as compared with the differenced tree cover measure.

We may also be more convinced that the differential uptake of agroforestry in the programme area was responsible for the corresponding increases in asset accumulation if those who took up the prompted practices and tree germplasm gained more than those who did not—*all else being equal*. A key issue, of course, is that those who took up the agroforestry dimension of Vi's programme may differ in both observable and unobservable ways from those who did not relevant to our outcomes of interest. Consequently, we implemented three different econometric procedures in an attempt to countervail this potential bias. The results are presented in Table 4.11.2.

To complement our mediation analysis, we also compared the effect estimates of high and low adoptions in the programme area, using three econometric modelling methods to control for bias.

The results of the first two sets of models presented in Table 4.11.2 compare high and low adopters in the programme area, respectively, with all households in the comparison area.[11] Here, we controlled for all the covariates in our dataset, save for baseline measures of the outcomes in question (see Annex 2). We further controlled for our other SALM and microenterprise participation indices, with fixed effects for our four VSZ dummy variables. As is clear, the coefficients for the high adopters are both statistically significant and significantly larger than those of the low adopters.

However, it is possible that there is one or more omitted variable that is correlated with either high and/or low adoption that we failed to (properly) measure and, hence, properly control for, which may account for the different effect estimates for these two groups. Consequently, we attempted to control for this potential unobserved bias using the Heckman Selection two-step estimator (Heckman 1978). In particular, we computed inverse Mills ratios to generate control functions for both the high adopters

[11] To be coded as a high adopter, the respondent's household had to have gained at 0.1 points on our 0-1 AF index and/or have gained an average increase of 10% tree cover on the plots of their main parcels.

and low adopters separately. We then included these control functions in the OLS regression models in an attempt to directly control for the unexplained (unobserved) variation in both high and low adoption. As is clear from Table 5.10.2, while the coefficients for the high adopters are downgraded slightly with the introduction of the control functions, there remain statistically significant and much larger than those of the low adopters.

TABLE 4.11.2: Associations between AF Adoption & Household Asset Wealth

	Overall			Female Respondent HHs		
	Consumption weighted dif. asset index	2016 PCA Asset Index	Asset Gain Index	Consumption weighted dif. asset index	2016 PCA Asset Index	Asset Gain Index
OLS (High adopters)						
Coefficient	0.20***	0.090***	0.11***	0.18**	0.13***	0.13***
	(0.068)	(0.029)	(0.028)	(0.081)	(0.035)	(0.034)
Observations	**2072**	**2072**	**2072**	**1256**	**1256**	**1256**
OLS (Low adopters)						
Coefficient	-0.0097	-0.00067	0.019	-0.0090	0.039	0.048
	(0.072)	(0.029)	(0.028)	(0.079)	(0.034)	(0.032)
Observations	**2096**	**2096**	**2096**	**1278**	**1278**	**1278**
OLS (High adopters with control functions)						
Coefficient	0.18***	0.087***	0.10***	0.16**	0.12***	0.13***
	(0.068)	(0.029)	(0.028)	(0.081)	(0.035)	(0.034)
Observations	**2066**	**2066**	**2066**	**1250**	**1250**	**1250**
OLS (Low adopters with control functions)						
Coefficient	-0.037	-0.0020	0.013	-0.036	0.033	0.040
	(0.072)	(0.029)	(0.028)	(0.077)	(0.033)	(0.031)
Observations	**2096**	**2096**	**2096**	**1278**	**1278**	**1278**
2SLS (Binary Adoption)						
Coefficient	0.92	0.47*	0.64**	0.83	0.81***	0.89***
	(0.58)	(0.25)	(0.26)	(0.65)	(0.31)	(0.31)
Observations	**2784**	**2784**	**2784**	**1702**	**1702**	**1702**
2SLS (AF Index)						
Coefficient	1.77*	0.87*	1.18***	1.64	1.49***	1.62***
	(1.06)	(0.45)	(0.44)	(1.20)	(0.52)	(0.51)
Observations	**2790**	**2790**	**2790**	**1708**	**1708**	**1708**
2SLS (Tree cover change)						
Coefficient	0.033	0.017*	0.023**	0.041	0.040**	0.044**
	(0.022)	(0.0093)	(0.0096)	(0.033)	(0.019)	(0.020)
Observations	**2784**	**2784**	**2784**	**1702**	**1702**	**1702**

Standard errors in parentheses and clustered at farmer group level; * p<0.1, ** p<0.05, *** p<0.01
All covariates in dataset used in OLS models; covariates correlated with Programme Area used in 2SLS models

The final three results pertain to our implementation of two-stage least square (2SLS) regression. However, this time, instead of instrumenting our Vi group dummy variable against programme area, we instrumented the following measures in each of the three models, respectively: the above binary measure of adoption, our AF Index, and our remote sensing derived differenced measure of tree cover change. In theory, these local average treatment effect (LATE) estimates apply specifically to those who were induced to become high agroforestry adopters via the coming of Vi to the programme area (Angrist and Pischke 2008). Here, we need to assume that the introduction of Vi's programme only positively induced households in the programme area into being significant agroforestry adopters, while not inducing others to move from high to low adopter status (i.e. the monotonicity assumption) (Blundell and Costa Dias 2009). This estimator works by taking the probability of being a high adopter whether one is located in the programme or comparison area net of their observable characteristics, and then regressing this probability on the outcome variable in

question when controlling for the same observable characteristics and fixed effects for VSZ.

Because the probability of being a high adopter is only around 11% higher in the programme area vis-à-vis the comparison area, the original OLS effect estimates are scaled up considerably.[12] As is clear from Table 4.11.2, the 2SLS models for high agroforestry adoption generated effect estimates that are considerably larger than those estimated for the high adopters. Nevertheless, the resulting high standard errors, renders several of these estimates statistically insignificant. It is further worth pointing out that some scholars have questioned what such LATE estimates actually identify and how they should be interpreted (Deaton 2010).

Hypothesis 2: The increased uptake of agroforestry led to increased income from the sale of agroforestry products and/or increased milk yields among dairy farmers, which, in turn led to greater asset accumulation

Our theory of change for Vi's programme posits several pathways through which the uptake of agroforestry was expected to lead to its expected final outcomes and impacts, including increased asset accumulation. Here, we first explore whether this could have been through increased income from the sale of agroforestry products and/or increased milk yields (presumably leading to increased milk-related income) among dairy farmers. To do so, we undertook mediation analysis using the following pathway sequence: programme area → mediator 1 → mediator 2 → asset outcome measure.

Table 4.11.3: Results of Mediation Analysis—Programme Effects on Asset Measures via AF Index + AF Sales

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Total Effect	**Direct Effect**	**Indirect Effect (Overall)**	**PA + AF index→ sales**	**p̂ mediated total**	**p̂ mediated AF index**	**p̂ mediated AF sales**	**p̂ mediated AF index + sales**
Overall								
Differenced Consump. Weighted Index	0.108* (.059)	0.60 (.059)	0.048*** (0.012)	0.286*** (0.04)	0.446	0.382	0.036	0.028
PCA 2016 Asset Index	0.048* (0.025)	0.031 (0.026)	0.017*** (0.005)	0.286*** (0.04)	0.358	0.345	0.008	0.006
Differenced Asset Gain Index	0.067*** (0.025)	0.055** (0.025)	0.013*** (0.005)	0.286*** (0.04)	0.189	0.217	-0.022	-0.017
Distressed Asset Sale	-0.047** (0.018)	-0.044** (0.018)	-0.003 (0.003)	0.273*** (0.05)	0.061	0.022	0.023	0.016
HHs with Female Participants Only								
Differenced Consump. Weighted Index	0.076 (0.068)	0.026 (0.069)	0.051*** (0.016)	0.234*** (0.039)	0.662	0.434	0.136	0.092
PCA 2016 Asset Index	0.080*** (0.03)	0.056* (0.031)	0.024*** (0.007)	0.235*** (0.039)	0.295	0.176	0.071	0.048
Differenced Asset Gain Index	0.089*** (0.029)	0.071*** (0.029)	0.019*** (0.007)	0.235*** (0.039)	0.209	0.114	0.056	0.038
Distressed Asset Sale	-0.074*** (0.027)	-0.060** (0.027)	-0.014*** (0.004)	0.219*** (0.045)	0.187	0.154	0.021	0.012

Standard errors in parentheses and clustered at the farmer group level; * p<0.1, ** p<0.05, *** p<0.01
Mediation analysis models include covariates correlated programme area and the various with mediator variable measures.
Mediation models all saturated, i.e. they were constructed to perfectly reproduce all variances, co-variances, and means.

[12] Despite this, our post estimation test of reveals that programme area in relation to high agroforestry adoption is not a weak instrument: F = 2.025 versus minimal critical value for 2SLS Size of nominal 5% Wald test of 16.38 (10%).

While there is evidence that the programme positively affected agroforestry product income and milk yields, there is little evidence that either was responsible for its estimated effects on asset accumulation.

Table 4.11.3 presents the results to assess the extent to which the programme's estimated effects on asset accumulation were mediated by the AF index[13] and agroforestry product sales (placed on a logarithmic scale). While similar, the Total Effect, Direct Effect, and overall Indirect Effect coefficients (columns 2, 3, and 4, respectively) slightly differ from those presented in Table 4.11.1, given that the exclusion of the differenced tree cover measure and inclusion of the differenced asset sales measure. The Indirect Effect estimates presented in column 5 indicate the extent to which the variation in data shared between programme area and the AF index explain variation in agroforestry product sales.

As is clear, this shared variation accounts for a significant amount, given that the indirect effect estimates are all statistically significant. However, in columns 6-9, we present the proportion of the effects of programme area that are mediated in total and by the AF index on its own, the AF product sale measure on its own, and the variation shared between the AF index and the AF product sale measure. As is clear, the vast share of the total mediated effect pertains to the independent variation explained by the AF index; the proportions mediated by AF sales independently and the AF index and sales combined are very modest. It follows, therefore, that there is little evidence that the pathway through which the increased uptake of agroforestry led to increased asset accumulation was via increased income through the sale of agroforestry products.

Table 4.11.4: Mediation Analysis—Programme Effects on Asset Gain Index via AF Index/Fodder & Milk Yields

(1)	(2) Total Effect	(3) Direct Effect	(4) Indirect Effect (Overall)	(5) PA + AF index/ fodder→ milk y.	(6) $\hat{p}$ mediated total	(7) $\hat{p}$ mediated AF /fodder	(8) $\hat{p}$ mediated milk yields	(9) $\hat{p}$ mediated AF/fod. + milk yields
General AF Index Pathway								
Differenced Asset Gain Index	0.077** (0.035)	0.074** (0.036)	0.003 (0.004)	0.012** (0.007)	0.043	0.019	0.021	0.002
Tree Fodder Pathway								
Differenced Asset Gain Index	0.077** (0.035)	0.073** (0.036)	0.004 (0.004)	0.031*** (0.01)	0.056	0.033	0.017	0.006

Standard errors in parentheses and clustered at the farmer group level; * p<0.1, ** p<0.05, *** p<0.01
Mediation analysis models include covariates correlated programme area and the various with mediator variable measures.
Mediation models are all saturated, i.e. they were constructed to perfectly reproduce all variances, co-variances, and means.

A related potential pathway is relevant specifically for dairy farmers. Their uptake of agroforestry, specifically through the uptake of dairy fodder, may have bolstered milk yields, resulting in increased income and, in turn, greater asset accumulation. For dairy farmers, the only statistically significant programme effect estimates among the asset accumulation measures pertain to the differenced asset gain index. Thus, we focused our mediation analysis only on this outcome measure, the results of which are presented in Table 4.11.4. As is clear, the differences between the Total Effect (column 2) and the Direct Effect (column 3) are minor, so it is not surprising that the two Indirect Effect estimates (column 4) are small and statistically insignificant. Consequently, the proportion of the programme's effect that could have been mediated is very small, with very little being mediated by the shared variation between the AF index or tree fodder use, on the one hand, and changes in milk yields on the other. However and consistent with the analysis of tree fodder use presented

[13] We focus on the AF index, rather than estimated changes in plot-level tree cover, given that the proportion of the programme's estimated effects that are mediated through it are significantly greater.

in Subsection 4.7, the shared variation between programme area and the AF Index and, more specifically, tree fodder uptake (column 5) significantly explains positive changes in milk yields

Hypothesis 3: The increased uptake of agroforestry improved soil health, which bolstered agricultural production and/or profits, which, in turn, led to greater asset accumulation

The soil health improvement hypothesis can be ruled out because our soil health measures and asset measures are not appropriately associated.

Our theory of change for Vi's programme also assumes that the uptake of agroforestry improved soil health, which ultimately translated into enhanced household asset accumulation via its positive effects on bolstering agricultural production. Unfortunately, we did not collect quantitative data on agricultural productivity, but we can assess the extent to which the data support the assertion that the programme's effects on asset accumulation were mediated through improvements in soil health. As we have seen, statistical associations between programme area and the mediator variable on the one hand and the mediator variable and the outcome variable on the other are a prerequisite for mediation to take place. For our soil organic carbon measure at least, the former criterion is met, while for soil erosion, this is less clear.

TABLE 4.11.5: Associations between Soil Organic Carbon & Erosion Measures & Asset Measures

	Overall	Sirisia/ Malakisi VSZ	Bumula VSZ	Kimilili/ Ndivisi VSZ	Tongaren/ Likuyani VSZ	Baseline Asset Rich	Female Part. HHs
Soil Organic Carbon							
Consumption weighted differenced asset index (OLS)							
Coefficient	-0.00094	0.00074	-0.034***	0.0083	-0.00047	-0.0024	0.0013
	(0.0043)	(0.0080)	(0.011)	(0.0098)	(0.0075)	(0.0067)	(0.0052)
Observations	**2785**	**690**	**717**	**676**	**702**	**1393**	**1707**
Asset gain index, PCA (OLS)							
Coefficient	-0.00084	-0.00061	-0.010*	-0.0015	0.00044	-0.0020	0.000094
	(0.0016)	(0.0032)	(0.0053)	(0.0034)	(0.0026)	(0.0025)	(0.0019)
Observations	**2785**	**690**	**717**	**676**	**702**	**1393**	**1707**
Distressed Asset Sale							
Coefficient	0.0100	0.022	0.065**	0.025	0.015	0.0025	-0.0041
	(0.0064)	(0.016)	(0.026)	(0.022)	(0.013)	(0.0084)	(0.0072)
Observations	**2015**	**503**	**509**	**369**	**422**	**970**	**1182**
Soil Erosion							
Consumption weighted differenced asset index (OLS)							
Coefficient	0.17	0.50	-0.020	0.31	0.31	0.15	0.079
	(0.17)	(0.34)	(0.37)	(0.32)	(0.56)	(0.25)	(0.18)
Observations	**2785**	**690**	**717**	**676**	**702**	**1393**	**1707**
Asset gain index, PCA (OLS)							
Coefficient	0.096	0.35**	0.018	0.064	0.031	0.16	0.10
	(0.073)	(0.14)	(0.17)	(0.13)	(0.20)	(0.11)	(0.078)
Observations	**2785**	**690**	**717**	**676**	**702**	**1393**	**1707**
Distressed Asset Sale							
Coefficient	-0.78**	-0.35	-1.30*	-0.40	-0.87	0.0081	-0.28
	(0.38)	(0.75)	(0.69)	(1.02)	(1.09)	(0.45)	(0.34)
Observations	**2015**	**503**	**509**	**369**	**422**	**970**	**1182**

Robust standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01
All covariates associate with dataset, save for baseline measures of the outcome variables, included in all models.

Table 4.11.5 presents results where we assessed the extent to which the latter criterion is met. As is clear, for both the soil organic carbon and soil erosion measures, all the overall coefficients are statistically insignificant, save for that for the distressed asset sales binary outcome for soil erosion. Here, those with greater soil erosion are less likely to have reported distressed asset sale. This not only does not support our

theory of change but also runs contrary to it. Moreover, at the VSZ level, there are several statistically significant associations, but these also are in an undesirable direction. Thus, even without carrying out formal mediation analysis, we can rule out the hypothesis that the programme's estimated effects on asset accumulation were due to the programme's positive effects on soil health, most particularly soil organic carbon.

Hypothesis 4: The increased uptake of agroforestry improved access to firewood, which increased income directly and indirectly, which, in turn led to greater asset accumulation

Our final hypothesis for how the programme's estimated impact on asset accumulation may have come about is perhaps less obvious but still worth exploring. This is via its estimated effect—presented in Subsection 4.6—on increasing the on-farm availability of firewood. It is possible that having more readily accessible firewood, for example, could have freed up women's time, in particular, for engaging in other productive pursuits, thereby promoting greater asset accumulation and/or reduced household expenditure on fuelwood.

Our fuel wood access measures are also only marginally associated with our asset accumulation measures, thereby ruling out the candidacy of increased fuelwood access as a key impact bridging mechanism.

As is the case for the soil health measures, Table 4.11.6 presents the results of OLS models that assessed the extent to which the firewood cash value and firewood collection time intermediary outcome measures are associated with the asset related asset measures. As is clear, the former intermediary outcome measure is only associated with the distressed asset sale variable, and the direction is the wrong way; households that reported gains in the cash value of their on-farm firewood are also more likely to report distressed asset sales.

TABLE 4.11.6: Associations between Firewood Cash Value & Collection Time & Asset Measures

	Overall	Sirisia/ Malakisi	Bumula	Kimilili/ Ndivisi	Tongaren/ Likuyani	Baseline Asset Rich	Female Part. HHs
Changes in Firewood Cash Value Collected on Farm							
Consumption weighted differenced asset index (OLS)							
Coefficient	0.011	0.0056	0.017	0.014	0.0083	0.0067	0.0052
	(0.0088)	(0.017)	(0.016)	(0.017)	(0.019)	(0.015)	(0.011)
Observations	**2785**	**690**	**717**	**676**	**702**	**1393**	**1707**
Asset gain index, PCA (OLS)							
Coefficient	0.0050	0.0057	0.00096	0.0076	0.0060	0.0050	0.0031
	(0.0036)	(0.0082)	(0.0069)	(0.0069)	(0.0068)	(0.0059)	(0.0041)
Observations	**2785**	**690**	**717**	**676**	**702**	**1393**	**1707**
Distressed Asset Sale							
Coefficient	0.028**	0.065**	0.051*	0.046	0.0059	0.024	0.015
	(0.013)	(0.028)	(0.027)	(0.039)	(0.029)	(0.019)	(0.015)
Observations	**2015**	**503**	**509**	**369**	**422**	**970**	**1182**
Changes in Time Spent Collecting Firewood (estimated hours per month)							
Consumption weighted differenced asset index (OLS)							
Coefficient	-0.032	-0.057	-0.087	-0.0076	0.057	-0.033	-0.016
	(0.030)	(0.066)	(0.054)	(0.072)	(0.060)	(0.049)	(0.037)
Observations	**2727**	**680**	**695**	**667**	**685**	**1358**	**1662**
Asset gain index, PCA (OLS)							
Coefficient	-0.017	-0.036	-0.060***	0.020	0.029	-0.020	-0.0044
	(0.012)	(0.028)	(0.023)	(0.026)	(0.024)	(0.021)	(0.014)
Observations	**2727**	**680**	**695**	**667**	**685**	**1358**	**1662**
Distressed Asset Sale							
Coefficient	0.15***	0.094	0.28***	0.045	0.16	0.20***	0.16***
	(0.051)	(0.097)	(0.096)	(0.23)	(0.12)	(0.074)	(0.060)
Observations	**1982**	**499**	**501**	**362**	**404**	**921**	**1161**

Roust standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01

All covariates associate with dataset, save for baseline measures of the outcome variables, included in all models.

However, the direction is desirable for the second intermediary outcome measure; increases in time spent collecting firewood are associated with increases in distressed asset sales and, by extension, the reverse for decreases in time collecting firewood. Further mediation analysis revealed that the overall indirect effect estimate is statistically significant. Nevertheless, this hypothesised mediator variable only accounts for approximately 5% of the programme area's total effect on distressed asset sale. Moreover, the indirect effect for both female participant households and those of Bumula VSZ are statistically insignificant (p=0.544 and p=0.277, respectively).

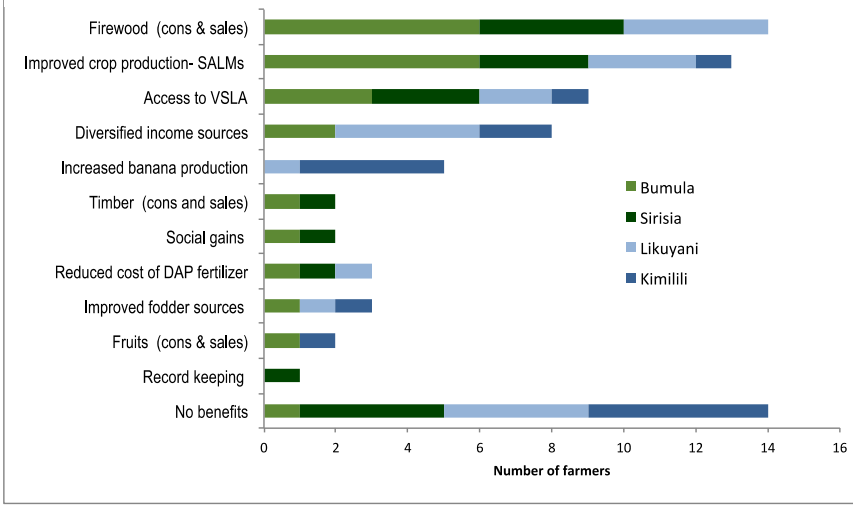
Qualitative insights on hypothesized mechanisms from the sub-sample of 40 Vi programme participants

Consistent with the quantitative results for the Economic Ladder question, 35 out of 40 of the Vi programme participants interviewed as part of our qualitative follow-up study stated that they are wealthier now than 10 years ago. And 23 stated that support from Vi had contributed to this. While two reported that Vi was solely responsible, 21 mentioned that other development actors also played a role, e.g. One Acre Fund.[14] When asked to list and rank the top three benefits of collaborating with Vi Agroforestry, 26 farmers mentioned at least one benefit, while 14 reported no benefits. From the latter, two had only been engaged in Vi's programme in a minor way, one had planted trees just recently, and nine farmers had never received any tree seeds/seedlings or training from Vi.

The greatest reported benefit from collaborating with Vi from the qualitative follow-up interviews was increased access to firewood, followed by improved crop production and access to credit through Village Savings and Loan Associations.

Figure 4.11.2 presents a breakdown of the reported benefits, with increased access to firewood and improved crop production through the adoption of Sustainable Agricultural Land Management (SALM) practices topping the list, followed by Village Savings and Loan Association (VSLA) participation and income diversification.

Figure 4.11.1: Ranked Benefits Reported from Vi Collaboration by Site



Increased access to firewood was reported to reduce the need to purchase it and time spent having to collect it. A woman farmer in Nabureyya in the Bumula VSZ, for example, explained that in the past, she used to rely on sugarcane and maize stalks

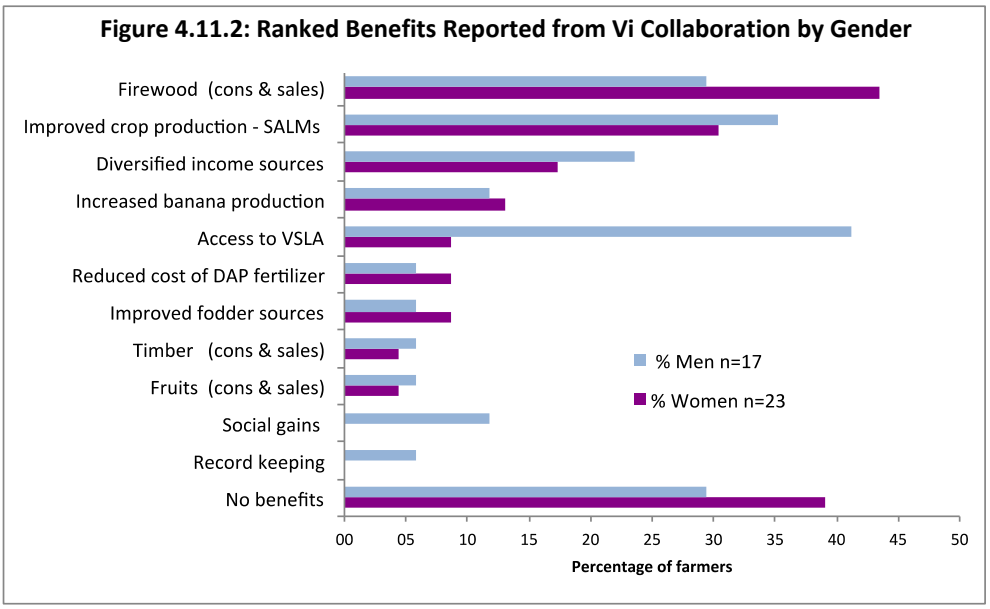
[14] Incidentally, we knew that One Acre Fund was active throughout the impact study area during the early stages of our study. Consequently, explicit measures were undertaken to ensure that the programme and comparison villages were evenly covered under its programmatic work.

but is now self-sufficient in firewood and sells over six tons a year from her four acre farm. Another Vi group member from Binyenya village in Sirisia mentioned she now has enough firewood for self-consumption; before 2014 she used to spend about USD $100 per year on firewood but now earns approximately $60 per year from selling it. For the other SALM practices reported by Vi, crop production was reported as having had increased through the uptake of composting and crop rotation. One woman participant from Sirisia reported that her production increased from three bags of maize to 10 bags in the last five years. Crop diversification was also mentioned as beneficial to overall production, as well as the benefits of participating in VSLAs, primarily via accessing loans for investing in income generation activities and to pay for school fees. For example, one participant from Sirisia mentioned he was able to buy a cow through a VSLA loan, while and another from Bumula opened a timber yard supplied by other members of his Vi affiliated farmer group, thereby creating synergy with Vi's promotion of tree planting and management activities.

Unfortunately, our qualitative data do not provide evidence of one or more clear mechanisms on how the greater uptake of agroforestry in the programme area actually translated into greater asset accumulation.

Unfortunately, the qualitative data do not provide clear evidence on the specific mechanisms in between agroforestry adoption and asset accumulation either. While increased access to firewood is reported as the top benefit, the quantitative data do not support this a being a critical intermediary mechanism. In fact, the qualitative data further forces us to re-examine two non-agroforestry related mechanisms that were earlier ruled out, i.e. the relatively greater increase in asset accumulation in the programme area may have resulted from a greater the uptake of other SALM practices and/or VSLA participation. We are confident in ruling out the former, given that we have good data on the actual uptake of the various SALM practices that were promoted and were able to use several quantitative tests to rule it out, as reported above. Our dismissal of a possible micro-credit/enterprise participation mechanism (particularly VSLA) through quantitative means is arguably weaker, as we did not go to considerable efforts to measure this comprehensively. However, the qualitative data makes up for this. Figure 4.11.2, in particular, presents a gender breakdown of ranked benefits of Vi collaboration. As is clear, very few women (two out of the 23 interviewed) reported access to VSLA as a key benefit. Given that the asset accumulation programme effect estimates primarily apply to households represented by female members, this makes VSLA participation an unlikely mechanism.

Figure 4.11.2: Ranked Benefits Reported from Vi Collaboration by Gender



The programme's estimated effects on asset accumulation are explained significantly by agroforestry adoption. However, the proceeding mechanisms for how such adoption translated into such effects is unclear.

In this Subsection we examined the extent to which four hypothesized mechanisms—derived from our theory of change for Vi's programme—that may account for Vi programme's estimated effects on asset accumulation—are consistent with the data collected through our study. We have seen that the first overarching hypothesised mechanism, i.e. that the relatively greater uptake of agroforestry was responsible is that which is most consistent. Indeed, a significant proportion of the programme's direct effect is explained by our measures of agroforestry uptake, most notably our AF Index. This is not the case for measures of the uptake of the other key dimensions of Vi's programme—i.e. other sustainable land management practices and micro-enterprise participation. In addition, significant asset accumulation effects are only particular to high agroforestry adopters, even following efforts to control for both observable and unobservable respondent and household characteristics. However, the bridging mechanism(s) from agroforestry uptake to asset accumulation is (are) unclear. The quantitative data do not, in particular, support the hypotheses that the relatively greater asset accumulation happened through correspondingly intermediary effects of greater agroforestry adoption on increased income from the sales of agroforestry products, improved soil health, and/or increased access to fuel wood. Our qualitative data does not shed much light on this either, unfortunately.

5. Discussion and Policy Implications

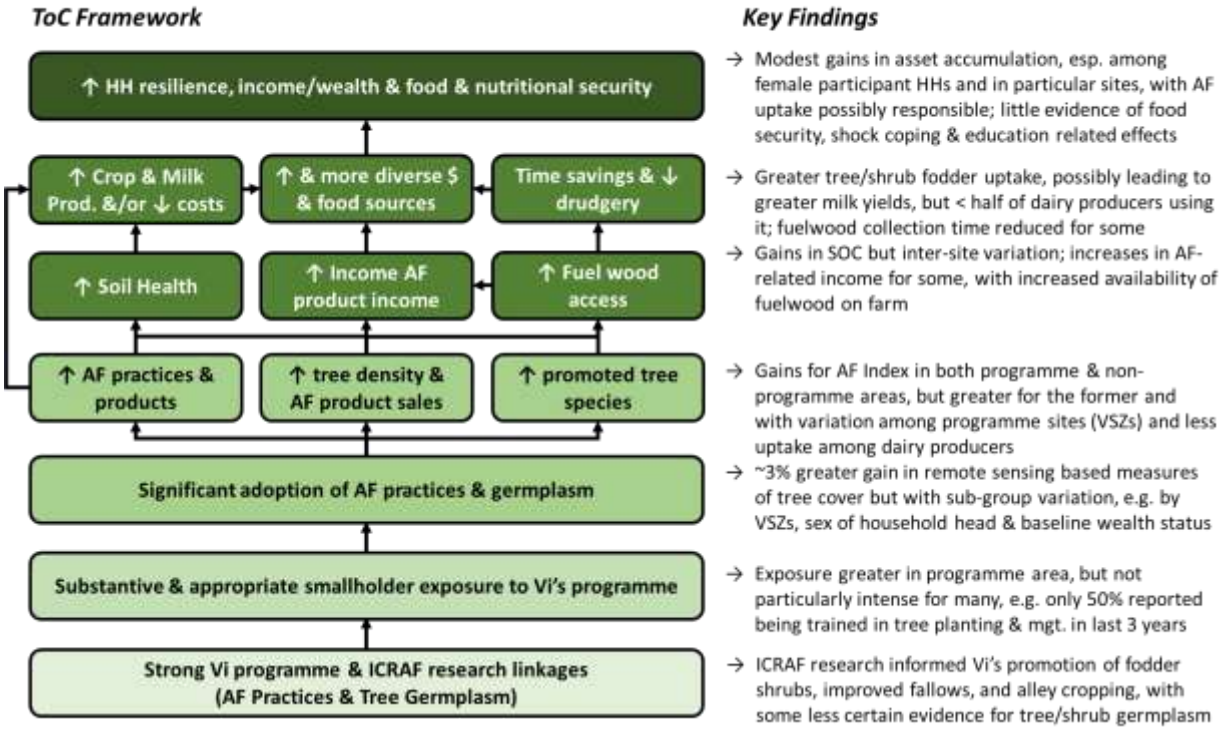
Evaluating the downstream impacts of agroforestry is intrinsically challenging.

The impact assessment on which this report is based is an ambitious attempt to address a key gap in the evidence base on the longer-term, more downstream effects of agroforestry promotion programmes and integrated agroforestry systems. This evidence gap has likely persisted to date simply because generating such rigorous evidence is particularly challenging when it comes to agroforestry: the timeframe within which such impacts are expected to manifest are typically long and non-linear; there is no one particular agroforestry system suitable for all agro-ecological, social, and economic contexts; agroforestry is generally taken up by farmers with varying levels of intensity, not as a binary 'technology'; and agroforestry promotion tends to be bundled with the promotion of other agricultural and natural resource management practices.

Bearing in mind these inherent challenges, we took advantage of an effort led by Vi Agroforestry to promote agroforestry in two counties in western Kenya, Bungoma and Kakamega. We chose these counties particularly because Vi had been operating within them for nine years and with the presence of adjacent geographic areas that could be used for comparison purposes where very little in the way of agroforestry promotion had taken place by either Vi itself or any other organization. Given that (a) no strong impact evaluation design was embedded into Vi's programme from the onset in general; and (b) this programme was rolled out to purposively selected—as opposed to randomly selected—areas, we had to go to considerable lengths to counter both programme placement and self-selection bias. We did this through (1) village-level matching using geospatial and secondary socioeconomic data; (2) sampling from all pre-existing farmer groups in the programme and comparison villages to estimate both intention to treat (ITT) effects and local average treatment effects (LATE); (3) reconstructing baseline data for difference-in-differences estimation; (4) using various econometric modelling approaches to control for measured differences between respondents and households in the programme and comparison areas; and (5) interrogating mechanisms via mediation analysis and rigorous qualitative approaches. Following our theory of change for Vi's programme, Figure 5.1 summarizes our main findings.

While we did find evidence of greater agroforestry programme exposure and, in turn, greater uptake of the particular and germplasm promoted by Vi in the programme area in general and among Vi groups in particular, neither can be considered as particularly profound, at least overall. In other words, there was significant variation in the extent of both programme exposure and agroforestry uptake. This was triangulated by both the quantitative and qualitative components of our study. It is, of course, difficult to know what would have happened if Vi's programme had been more intensely implemented across all the farmer groups and among all their respective members. However, one cannot help but speculate that, if agroforestry uptake had been significantly more intense, whether a different set of programme effect estimates would have been estimated. Indeed, while overtly conscious of the possibility of confounding, we saw in Subsection 4.11 that higher agroforestry adopters in the programme areas accumulated greater asset wealth than their lower adopting counterparts. Recall that this particular evaluation focused on Vi's post-2004 implementing model, which is evidently lighter on extension and follow-up than was previously the case; the earlier, more extension heavy model was not put to the test.

FIGURE 5.1: Summary of Key Findings along ToC



This connects this impact assessment to the body of evaluation literature on implementation fidelity (Wilson et al. 2010; Blakely et al. 2002; Durlak & DuPre 2008). In particular and perhaps as is intuitive, there is evidence from many studies that 'implementation matters' in social programmes; that is, provided that an intervention or programme is actually effective, those receiving more complete and/or better quality implementation are more likely to be positively affected. While infidelity to a programme's implementation model can be positive, i.e. associated with adaptation or innovation (Larsen and Samdal 2007), it is often simply the result of failing to put in place the requite conditions to ensure that appropriate participant exposure to the programme takes place (Williams 2007). Given that one of Vi's main modus operandi involves training farmer groups and only about half reported having had been trained in tree planting and management at least once in the past three years, it is clear that many members of the groups that Vi targeted were simply not well exposed to its programme.

Indeed, it is useful to view the promotion of agroforestry—as well as many other agricultural and natural resource management practices for that matter—as complex interventions. Borrowing from evaluative efforts undertaken in the public health sector, complex interventions are characterized as having one or more of the following attributes: (a) several interacting components; (b) significant and specific behaviour [change] required by those delivering and/or receiving the intervention; (c) more than one group or organizational level targeted by the intervention; (d) multiple potential outcomes and expected heterogeneity of those outcomes; and (e) need for tailoring or adapting the intervention to varying circumstances (Craig et al. 2008). It is clear that agroforestry promotion in general and Vi's programme in particular shares most, if not all, of these attributes. For obvious reasons, successfully implementing complex interventions, so that they have the potential of bringing about their expected results, requires significant attention to process, monitoring, review, and

It is useful to view the promotion of agroforestry as a complex intervention, which require considerable monitoring, ongoing follow-up support, and/or adaptation to ensure successful and impactful uptake.

adaptive management. Yet and generally speaking, this does not seem to have been the case for many participants targeted by Vi's programme. It appears that Vi may have moved too far away from its pre-2004 extension heavy implementation approach.

Another point worth mentioning pertains to agroforestry's general long-term and non-linear impact pathways. Arguably, many of the income generation benefits from Vi's programme would be expected from the sale of agroforestry products, such as timber and firewood. However, as we saw in Subsection 4.6, more than half of Vi programme participants reported no revenue from such sales over the last 12 months at all. In other words, the timing of our assessment may have been too soon for many programme participants, given that the full impacts of their newly established or scaled-up agroforestry systems had yet to fully manifest. This leads us to Woolcock's (2009) impact trajectory concept. He argues that obtaining reliable feedback on the effectiveness of development interventions does not solely rely on the identification of a plausible counterfactual. It also depends on, "…an appropriate match between the shape of impact trajectory over time and the deployment of a corresponding array of research tools capable of empirically discerning such a trajectory." And for the promotion of agroforestry across heterogeneous contexts, this adds an additional layer of complexity, given that the shape of such trajectories may be inherently different for different farming systems.

We conclude this report by making three specific recommendations, relevant for donors, development and research organizations, and/or policy makers promoting and researching agroforestry and other complex agricultural and natural resource management interventions and systems more broadly:

1. Build in explicit provisions for programme exposure and uptake monitoring, the provision of demand driven, follow-up support, and adaptive management when pursuing complex interventions, such as the promotion of agroforestry, with a view to ensuring comprehensive and appropriate programme exposure.

2. Set agroforestry promotional programmes/interventions up for rigorous evaluation from the onset, but experiment with innovative tools to estimate longer-term impacts and returns on investment.

3. Foster collaborative and synergy between researchers and practitioners to leverage greater delivery impact delivery among both.

References

Abadie, Alberto. 2005. “Semiparametric Difference-in-Differences Estimators.” *Review of Economic Studies* 72: 1–19.

Abadie, B Y Alberto, and Guido W Imbens. 2016. “Large Sample Properties of Matching Estimators for Average Treatment Effects Author (S): Alberto Abadie and Guido W . Imbens Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/3598929 Accessed : 04-06-2016 09 : 55 UTC Your U” 74 (1): 235–67.

Alkire, Sabina, and James Foster. 2011. “Counting and Multidimensional Poverty Measurement.” *Journal of Public Economics* 95 (7–8): 476–87. doi:https://doi.org/10.1016/j.jpubeco.2010.11.006.

Andam, Kwaw S, Paul J Ferraro, Alexander Pfaff, G Arturo Sanchez-Azofeifa, and Juan a Robalino. 2008. “Measuring the Effectiveness of Protected Area Networks in Reducing Deforestation.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (42): 16089–94. doi:10.1073/pnas.0800437105.

Andam, Kwaw S, Paul J Ferraro, Katharine R E Sims, Andrew Healy, and Margaret B Holland. 2010. “Protected Areas Reduced Poverty in Costa Rica and Thailand.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (22): 9996–10001. doi:10.1073/pnas.0914177107.

Angrist, Joshua, and Jorn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton University Press. citeulike-article-id:4423860.

Bilinsky, Paula, and Anne Swindale. 2010. “Months of Adequate Household Food Provisioning (MAHFP) for Measurement of Household Food Access: Indicator Guide (v.4).” Washington, D.C.

Blakely, Craig H, Jeffrey P Mayer, Rand G Gottschalk, Neal Schmitt, William S Davidson, David B Roitman, and James G Emshoff. 2002. “The Fidelity-Adaptation Debate: Implications for the Implementation of Public Sector Social Programs BT - A Quarter Century of Community Psychology: Readings from the American Journal of Community Psychology.” In , edited by Tracey A Revenson, Anthony R D’Augelli, Sabine E French, Diane L Hughes, David Livert, Edward Seidman, Marybeth Shinn, and Hirokazu Yoshikawa, 163–79. Boston, MA: Springer US. doi:10.1007/978-1-4419-8646-7_10.

Bland, J Martin, and Douglas G Altman. 1997. “Statistics Notes: Cronbach's Alpha.” *BMJ* 314 (7080): 572. http://www.bmj.com/content/314/7080/572.abstract.

Blundell, Richard, and Monica Costa Dias. 2009. “Alternative Approaches to Evaluation in Empirical Microeconomics.” *Journal of Human Resources* 44 (3): 565–640. doi:10.1353/jhr.2009.0009.

Buttoud, G., F. Place, and M. Gauthier. 2013. “Advancing Agroforestry on the Policy Agenda.” *Agroforestry Working Paper No.1*. Vol. FAO, Rome,. doi:10.1080/14728028.2013.806162.

Collier, David. 2011. “Understanding Process Tracing.” *Political Science and Politics* 44 (4): 823–30. doi:10.1017/S1049096511001429.

Craig, Peter, Paul Dieppe, Sally Macintyre, Susan Michie, Irwin Nazareth, and Mark Petticrew. 2008. “Developing and Evaluating Complex Interventions: The New Medical Research Council Guidance.” *BMJ* 337 (September). http://www.bmj.com/content/337/bmj.a1655.abstract.

Deaton, Angus. 2010. “Instruments, Randomization, and Learning about Development.” *Journal of Economic Literature* 48 (2): 424–55. http://www.aeaweb.org/articles?id=10.1257/jel.48.2.424.

Deaton, Angus, and Salman Zaidi. 1999. “Guidelines for Constructing Consumption Aggregates For

Welfare Analysis." 217.

Douthwaite, B., V. M. Manyong, J. D H Keatinge, and J. Chianu. 2002. "The Adoption of Alley Farming and Mucuna: Lessons for Research, Development and Extension." *Agroforestry Systems* 56 (3): 193–202. doi:10.1023/A:1021319028117.

Durlak & DuPre. 2008. "Implementation Matter : A Review of Research on The Influence of Implementation on Program Outcomes and The Factors Affecting Implementation," 327–50.

FAO, and FHI 360. 2016. "Minimum Dietary Diversity for Women: A Guide for Measurement." Rome.

Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, et al. 2015. "The Climate Hazards Infrared Precipitation with Stations: A New Environmental Record for Monitoring Extremes." *Scientific Data* 2. doi:10.1038/sdata.2015.66.

Funk, Michele Jonsson, Daniel Westreich, Chris Wiesen, Til Stürmer, M. Alan Brookhart, and Marie Davidian. 2011. "Doubly Robust Estimation of Causal Effects." *American Journal of Epidemiology* 173 (7): 761–67. doi:10.1093/aje/kwq439.

Handa, Sudhanshu, and John A. Maluccio. 2010. "Matching the Gold Standard: Comparing Experimental and Nonexperimental Evaluation Techniques for a Geographically Targeted Program." *Economic Development and Cultural Change* 58 (3): 415–47. doi:10.1086/650421.

Heckman, James J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46 (4). [Wiley, Econometric Society]: 931–59. doi:10.2307/1909757.

Honey-Rosés, Jordi, Kathy Baylis, and M. Isabel Ramírez. 2011. "A Spatially Explicit Estimate of Avoided Forest Loss." *Conservation Biology* 25 (5): 1032–43. doi:10.1111/j.1523-1739.2011.01729.x.

Hughes, Karl Alan. 2012. "Where There Is No External Control : Exploring the Use of Treatment Exposure Variation and Programme Theory to Manage for Outcome Success (PhD Thesis)." London School of Hygeine and Tropical Medicine.

Imbens, Guido W. 2010. "Better LATE Than Nothing - Comments." *Journal of Economic Literature* 48 (June): 399–423.

Jarvis, A, HI Reuter, A Nelson, and E Guevara. 2008. "Hole-Filled SRTM for the Globe Version 4, Available from the CGIAR-CSI SRTM 90m Database."

Joppa, Lucas, and Alexander Pfaff. 2010. "Reassessing the Forest Impacts of Protection: The Challenge of Nonrandom Location and a Corrective Method." *Annals of the New York Academy of Sciences* 1185: 135–49. doi:10.1111/j.1749-6632.2009.05162.x.

Kruska, Russell, and Patrick Chege Kariuki. n.d. "ILRI - GIS Services Digital Elevation Model."

Larsen, Torill, and Oddrun Samdal. 2007. "Implementing Second Step: Balancing Fidelity and Program Adaptation." *Journal of Educational and Psychological Consultation* 17 (1). Routledge: 1–29. doi:10.1080/10474410709336588.

LePage Morgan, Claire. 2017. "From the Lab to the Farm: Using Process-Tracing Methods to Evaluate Impact Pathways in Agroforestry Research."

Leuven, E., and B. Sianesi. 2003. "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing."

MacKinnon, David Peter. 2008. *Introduction to Statistical Mediation Analysis*. New York: Lawrence Erlbaum Associates.

Maxwell, Daniel G., and Richard Caldwell. 2008. “The Coping Strategies Index Field Methods Manual.”

Morgan, Seth. 2017. “Where Trees Grow, Assets Grow: Applying Spatial Matching to Evaluate Agroforestry’s Household Welfare Impacts in Kenya.” University of Illinois at Urbana-Champaign.

Morris, S. S, Calogero Carletto, John Hoddinott, and L J M Christiaensen. 2000. “Validity of Rapid Estimates of Household Wealth and Income for Health Surveys in Rural Africa.” *Journal of Epidemiology & Community Health* 54 (5): 381–87. doi:10.1136/jech.54.5.381.

Moser, Caroline, and Andrew Felton. 2007. “The Construction of an Asset Index Measuring Asset Accumulation in Ecuador Caroline Moser and Andrew Felton CPRC Working Paper 87 ISBN 1-904049-86-9.” *Development*. http://www.mendeley.com/research/construction-asset-index-measuring-asset-accumulation-ecuador-caroline-moser-andrew-felton-cprc-working-paper-87-isbn-1904049869/.

Nair, P.K. Ramachandran. 1993. *An Introduction to Agroforestry*. *Forest Ecology and Management*. Vol. 73. DORDRECHT/ BOSTON / LONDON: Kluwer Academic Publishers. doi:10.1016/0378-1127(95)90008-X.

O’Donnell, Owen, Eddy van Doorslaer, Adam Wagstaff, and Magnus Lindelow. 2008. *Analyzing Health Equity Using Household Survey Data : A Guide to Techniques and Their Implementation*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/6896 License: CC BY 3.0 IGO.

Odhiambo, H. O., C. K. Ong, J. D. Deans, J. Wilson, a. a H Khan, and J. I. Sprent. 2001. “Roots, Soil Water and Crop Yield: Tree Crop Interactions in a Semi-Arid Agroforestry System in Kenya.” *Plant and Soil* 235 (2): 221–33. doi:10.1023/A:1011959805622.

Otsuki, Tsunehiro. 2010. “Estimating Agroforestry ’ S Effect on Productivity in Kenya : An Application of a Treatment Effects Model.” *Africa*, no. C: 2007–9.

Place, F, M Adato, P Hebinck, and M Omosa. 2005. “The Impact of Agroforestry-Based Soil Fertility Replenishment Practices on the Poor in Western Kenya.” *Research Report of the International Food Policy Research Institute*. http://www.scopus.com/inward/record.url?eid=2-s2.0-33947281535&partnerID=40&md5=96562b0ae8a5fd22cb2280a04495a23a.

Punton, Melanie, and Katharina Welle. 2015. “Straws-in-the-Wind, Hoops and Smoking Guns: What Can Process Tracing Offer to Impact Evaluation ?” CDI Practice Paper.

Ravallion, Martin, and Michael Lokshin. 1999. “Subjective Economic Welfare.” Policy, Research Working Paper ; No. WPS 2106. Washington, DC. http://documents.worldbank.org/curated/en/996191468757792826/Subjective-economic-welfare.

Robalino, Juan, Alexander Pfaff, and Laura Villalobos. 2015. “Deforestation Spillovers from Costa Rican Protected Areas.” *Documentos de Trabajo U. Costa Rica*.

Rosenbaum, Paul R, and Donald B Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70 (1083): 41–55.

Sexton, J.O., X.P. Song, M. Feng, P Noojipady, A. Anand, C. Huang, D.H. Kim, et al. 2013. “Landsat Tree Cover Continuous Fields.” doi:10.1080/17538947.2013.786146.

Sileshi, Gudeta, Festus K Akinnifesi, Oluyede C Ajayi, and Frank Place. 2009. “Evidence for Impact of Green Fertilizers on Maize Production in Sub-Saharan Africa: A Meta-Analysis.”

Sjögren, Hans, Keith D. Shepherd, and Anders Karlsson. 2010. “Effects of Improved Fallow with Sesbania Sesban on Maize Productivity and Striga Hermonthica Infestation in Western Kenya.” *Journal of Forestry Research* 21 (3): 379–86. doi:10.1007/s11676-010-0085-0.

Stevens, Forrest R., Andrea E. Gaughan, Catherine Linard, and Andrew J. Tatem. 2015. “Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data.” *PLoS ONE* 10 (2): 1–22. doi:10.1371/journal.pone.0107042.

Vågen, Tor-G., Leigh A Winowiecki, Jerome E Tondoh, Lulseged T Desta, and Thomas Gumbricht. 2016. “Mapping of Soil Properties and Land Degradation Risk in Africa Using MODIS Reflectance.” *Geoderma* 263. Elsevier B.V.: 216–25. doi:http://dx.doi.org/10.1016/j.geoderma.2015.06.023.

Vågen, Tor G., Leigh a. Winowiecki, Assefa Abegaz, and Kiros M. Hadgu. 2013. “Landsat-Based Approaches for Mapping of Land Degradation Prevalence and Soil Functional Properties in Ethiopia.” *Remote Sensing of Environment* 134 (August): 266–75. doi:10.1016/j.rse.2013.03.006.

Wachiye, Emmanuel. 2008. “AF AND SALMS Presentation - in Use since 2008.”

White, H. 2010. “A Contribution to Current Debates in Impact Evaluation.” *Evaluation* 16 (2): 153–64. doi:10.1177/1356389010361562.

Williams, Eddie. 2007. “Extensive Reading in Malawi: Inadequate Implementation or Inappropriate Innovation?” *Journal of Research in Reading* 30 (1). Blackwell Publishing Ltd: 59–79. doi:10.1111/j.1467-9817.2006.00328.x.

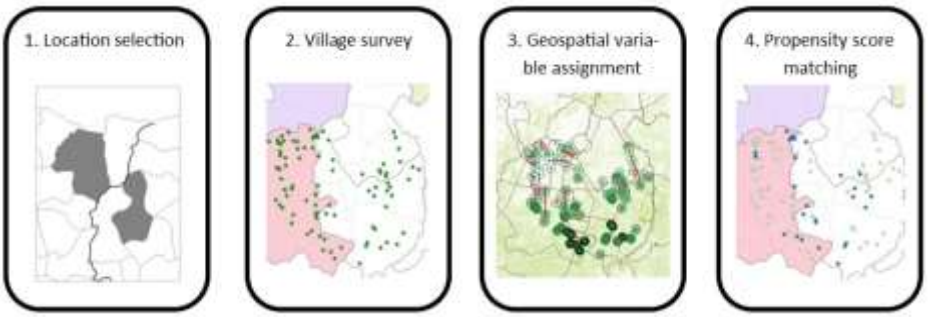
Wilson, Mark G., Tania B. Basta, Bethany H. Bynum, David M. Dejoy, Robert J. Vandenberg, and Rod K. Dishman. 2010. “Do Intervention Fidelity and Dose Influence Outcomes? Results from the Move to Improve Worksite Physical Activity Program.” *Health Education Research* 25 (2): 294–305. doi:10.1093/her/cyn065.

Woolcock, Michael. 2009. “Toward a Plurality of Methods in Project Evaluation: A Contextualised Approach to Understanding Impact Trajectories and Efficacy.” *Journal of Development Effectiveness* 1 (1): 1–14. doi:10.1080/19439340902727719.

Annexes

Annex 1: Village Matching using Spatial and Secondary Data

The execution of the programme area and comparison area village selection process followed a four-step procedure: (1) sub-location selection; (2) scoping survey administration; (3) geospatial variable assignment; and (4) propensity score matching.



Sub-location selection: Qualitative interviews were conducted to guide a purposive selection of sub-locations (the smallest administrative unit in Kenya). These sublocations were selected from Vi's program area and the non-program area based on their similarity in terms of relative wealth and agro-ecological characteristics. These interviews were carried out with Vi field staff, Kenya Ministry of Agriculture field officers, and farmer group leaders. Once these interviews were carried out, the sample was restricted to these purposively matched sublocations.

Scoping survey administration: A scoping survey was administered within the villages in the purposively matched sublocations. Local consultants were hired to visit the villages where they administered a short survey with key informants and captured the GPS coordinates of the centre of the village. The scoping survey instrument was implemented using the Open Data Kit (ODK) platform. It captured: (a) the number of households in the village; (b) whether there were active farmer groups which have been active from the baseline period onwards; and (c) the activities of each group and their receipt of NGO or government support, if any.

Geospatial variable assignment. The village GPS coordinates captured during the scoping survey were used to determine the nine geospatial variables listed below. A 1 km buffer was generated around each village's central geocode and the Zonal Statistics tool in ArcGIS was used to calculate the average value across the raster grids containing each variable's values. The village average values were then assigned to the dataset of village names using the Extract Multi Values by Points tool in ArcGIS.

Propensity Score Matching (PSM). The geospatial variables presented below were used to assign a propensity score, and the villages were matched on this score using *psmatch2*'s one-to-one calliper matching algorithm in Stata. Applying this method to villages allowed us to compare households in villages within the programme area to household in villages outside the programme area, which are similar across the range of relevant covariates. Matching assumes that average treatment effects can be estimated by taking the average of the difference between the expected outcome of untreated observations—conditional on a vector of covariates—from the expected outcome of the treated observations conditional on the same covariates (Abadie and Imbens 2016).

This methodology has been used to estimate the effects of protected areas (Honey-Rosés, Baylis, and Ramírez 2011; Andam et al. 2008, 2010). This literature includes examples of matching on observational units at multiple scales, including pixels in a raster grid (Robalino, Pfaff, and Villalobos 2015; Andam et al. 2008), polygons corresponding to land management units (Honey-Rosés, Baylis, and Ramírez 2011), and census tracts (Andam et al. 2010). Our analysis focuses on a 1 km circular buffer drawn around each village considered for the study.

Matching designs in the literature use two primary methodologies for assessing observations' similarity across covariates: nearest neighbour and propensity score matching (Joppa and Pfaff 2010). Nearest neighbour matching calculates the multi-dimensional distance between two observations given the vector of covariates. Propensity score matching condenses the covariates to a single score using a regression model to calculate each observation's conditional probability of receiving treatment given the covariate values (Rosenbaum and Rubin 1983). Propensity score matching was identified as the most appropriate method, since the objective was to identify comparison villages with a high conditional probability of being included in the programme area given the measured geospatial variables.

The propensity score is defined formally as the probability of treatment, conditional on a vector of covariates. The propensity score model can be expressed by the equation below:

$$e(X_i) = \Pr(T_i = 1 \mid X_i)$$

Where $e(X_i)$ is the probability of being included in the treatment group, and X_i is the vector of covariates listed above. The propensity score was generated using a probit model estimated within Stata by the psmatch2 command. The probit model takes the form:

$$z = X\beta + \varepsilon$$

Where z is an unobserved variable, X is a vector of covariates with coefficients β, and y is the observed binary corresponding to treatment assignment such that:

$$y_i = \begin{cases} 1 \; if \; z \geq 0 \\ 0 \; if \; z \leq 0 \end{cases}$$

Due to the limited time before enumeration was scheduled to begin, the scoping process was completed in four phases. In particular, the study area was split into four zones, which we refer to as Village Sampling Zones (VSZs) in this report, and the matching process was applied to each one separately as the scoping data were collected. The targeted number of villages in each VSZ was set as 15 programme area villages and 15 comparison area villages. Consequently, for each VSZ, the 30 best-matched villages were chosen by gradually reducing the calliper width using Stata's *psmatch2* command. Covariate balance was thereafter tested on the set of matching variables between the programme and comparison villages for each VSZ. Then, as the team of enumerators began collecting data in the matched villages in the VSZ in question, the scoping team moved forward to the next VSZ and the PSM village matching exercise was subsequently undertaken for this particular VSZ.

Villages with farmer-led groups active since the beginning of the study period (2008-2016) were considered candidates for inclusion in the study. Using geocodes from the centre of each villages, geospatial variables were assigned to each of these shortlisted villages for input into the matching model. The matching covariates included agro-ecological characteristics, as well as socio-economic indicators, such as population density and distance from major roads. We chose these particular covariates because we assume they are likely to affect agricultural production and market access, and, as such, they would likely be significant confounders of the measured treatment effect if the selected villages were unbalanced across them. The following variables, in particular, were used in the propensity score model:

- Number of Households
- Average Soil Sand Content (T.-G. Vågen et al. 2016)
- Average Soil pH (T.-G. Vågen et al. 2016)
- Average Soil Organic Carbon in 2007 (T.-G. Vågen et al. 2016)
- Average Tree Cover in 2005 (Sexton et al. 2013)
- Elevation (Jarvis et al. 2008; Kruska and Kariuki, n.d.)

- Average Population Density in 2010 and 2015 (Stevens et al. 2015)
- Average Rainfall (C. Funk et al. 2015)
- Distance to Tarmac Road
- Binary for Villages 0.25 m from Tarmac Road ("on road")
- Binary for presence of microfinance activities

Elevation, tree cover, population density, and soil variables were measured as an average value calculated across a circle 1 km in radius extending from a central point in the village. Rainfall was measured as the value of the raster cell in which the village centre was found. The cells for the Climate Hazards Infrared Precipitation with Stations (CHIRPS) rainfall dataset measure approximately 5.5 km across (C. Funk et al. 2015).

Household numbers were taken from the village-level scoping survey. The consultants requested the number of households from leaders of farmer groups, and, if they did not know this number, they requested it from a village elder. The tarmac road network was taken from OpenStreetMap data, and ground-truthed by travel in the region. The binary variable for the presence of microfinance activities was taken from Vi's records on their participating groups and from the scoping survey for the comparison villages. Activities listed as "table banking" or "merry-go-round" were counted as microfinance activities.

After each VSZ specific matching exercise, overall covariate balance was checked to confirm that the overall village sample of treatment and comparison villages are balanced across all selected covariates, as is presented in Table 1.

Table 1: Village-level Matching Balancing Statistics

	Sample Mean	Programme Mean	Non-Programme Mean	Normalized Difference	Difference
Soil Sand Content	19.96	20.57	19.36	0.11	1.21
					(1.42)
Soil pH	5.95	5.97	5.94	0.20	0.03
					(0.02)
Tree Cover 2005	6.07	5.97	6.17	-0.06	-0.21
					(0.45)
Elevation	1570.52	1575.63	1565.49	0.05	10.13
					(26.18)
Population Density 2010	4.41	4.40	4.43	-0.02	-0.03
					(0.22)
Soil Organic Carbon 2007	25.57	24.71	26.43	-0.19	-1.72
					(1.16)
Rainfall	136.71	133.97	139.40	-0.23	-5.42
					(2.90)
Distance to Road	0.03	0.03	0.03	0.07	0.00
					(0.00)
On Road	0.02	0.02	0.03	-0.07	-0.02
					(0.03)
Observations	**121**	**60**	**61**		

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses

Annex 2: Baseline and Time Invariant Characteristic Comparisons

A2.1: Comparison of HHs in Program and Non-program Areas—Binary Characteristics

Characteristic	Program Mean	Non-program Mean	Difference (raw)	Difference (net of county)	Difference (net of VSZ)
Respondent Female	0.63	0.60	0.031*	0.082*	0.082*
			(1.70)	(1.69)	(1.69)
Respondent Christian	0.99	0.99	-0.0027	-0.16	-0.17
			(-0.91)	(-0.92)	(-0.98)
Respondent Luhya	0.97	0.98	-0.0095	-0.14	-0.16
			(-1.49)	(-1.48)	(-1.64)
Respondent Married	0.75	0.73	0.017	0.050	0.054
			(1.01)	(0.98)	(1.06)
Res. Married by Spouse Elsewhere	0.08	0.08	0.0036	0.021	0.012
			(0.35)	(0.31)	(0.18)
Respondent Widowed	0.14	0.16	-0.018	-0.073	-0.077
			(-1.32)	(-1.25)	(-1.31)
Respondent is Head	0.48	0.54	-0.057***	-0.14***	-0.14***
			(-3.01)	(-3.02)	(-3.03)
Respondent is Spouse of Head	0.41	0.37	0.038**	0.099**	0.100**
			(2.06)	(2.06)	(2.07)
Respondent Good Health	0.97	0.98	-0.0089	-0.15	-0.16
			(-1.51)	(-1.50)	(-1.52)
Respondent Literate	0.89	0.89	0.00059	0.0039	0.0059
			(0.049)	(0.06)	(0.09)
Respondent in School	0.02	0.03	-0.0033	-0.057	-0.070
			(-0.55)	(-0.57)	(-0.69)
Respondent has Official Role	0.49	0.46	0.027	0.067	0.070
			(1.40)	(1.40)	(1.48)
Respondent Farmer (main occupation)	0.87	0.86	0.017	0.078	0.079
			(1.29)	(1.30)	(1.33)
Respondent Primarily Business	0.05	0.06	-0.0067	-0.061	-0.059
			(-0.77)	(-0.80)	(-0.77)
Respondent Employed	0.05	0.06	-0.013	-0.12	-0.13
			(-1.56)	(-1.56)	(-1.63)
Respondent has Tech. Skills	0.06	0.04	0.020**	0.18**	0.18**
			(2.30)	(2.31)	(2.29)
HH was farming in 2007	0.99	0.98	0.0046	0.13	0.12
			(0.99)	(1.02)	(1.01)
HH reared livestock in 2007	0.62	0.57	0.055***	0.14***	0.14***
			(2.97)	(2.97)	(2.92)
HH ran off-farm business in 2007	0.30	0.31	-0.00045	-0.0015	-0.0089
			(-0.026)	(-0.03)	(-0.18)
HH member employed in 2007	0.16	0.13	0.030**	0.13**	0.12**
			(2.24)	(2.21)	(2.08)
HH used irrigation in 2007	0.05	0.04	0.0077	0.078	0.086
			(0.96)	(0.96)	(1.04)
Head is productive	0.97	0.97	-0.0066	-0.10	-0.11
			(-1.04)	(-1.06)	(-1.11)
all adults in HH over 59	0.06	0.05	0.0047	0.041	0.042
			(0.54)	(0.53)	(0.54)
Head is 60 or older	0.33	0.29	0.043**	0.12**	0.12**
			(2.46)	(2.46)	(2.42)
female headed HH	0.22	0.22	-0.0040	-0.011	-0.015
			(-0.25)	(-0.20)	(-0.28)
Literate adult in HH	0.98	0.97	0.0054	0.095	0.095
			(0.93)	(0.93)	(0.93)
Female literate adult in HH	0.89	0.89	0.0026	0.013	0.011
			(0.22)	(0.20)	(0.17)
Head is literate	0.90	0.90	-0.0039	-0.022	-0.020
			(-0.34)	(-0.34)	(-0.30)
HH located on tarmac road	0.04	0.05	-0.010	-0.11	-0.11
			(-1.33)	(-1.32)	(-1.24)
HH had formal title to main parcel (07)	0.35	0.32	0.030*	0.082*	0.078
			(1.67)	(1.66)	(1.58)
Respondent owned main parcel (07)	0.46	0.54	-0.084***	-0.21***	-0.21***
			(-4.44)	(-4.44)	(-4.40)
Observations	**1411**	**1386**	**2797**	**2797**	**2797**

z statistics in parenthesis; VSZ=Village Sampling Zone; * p<0.1, ** p<0.05, *** p<0.01
Probit regression used for net of county and VSZ differences, so coefficients are not directly interpretable, only the *t*-statistics

A2.2: Comparison of HHs in Program and Non-program Areas—Continuous Characteristics

Characteristic	Program Mean	Non-program Mean	Difference (raw)	Difference (net of county)	Difference (net of VSZ)
Age of respondent	49.65	48.84	0.81	0.80	0.76
			(1.60)	(1.60)	(1.50)
Years of education of respondent	8.55	8.61	-0.055	-0.056	-0.061
			(-0.41)	(-0.41)	(-0.45)
Household size	6.53	6.65	-0.13	-0.13	-0.12
			(-1.36)	(-1.35)	(-1.30)
Number of children in HH	3.56	3.71	-0.15**	-0.15**	-0.15*
			(-1.99)	(-1.99)	(-1.91)
Number of adults in HH	2.97	2.94	0.025	0.025	0.024
			(0.46)	(0.46)	(0.44)
Years of education head	9.08	9.20	-0.12	-0.12	-0.12
			(-0.86)	(-0.86)	(-0.88)
Highest years of educ. of any adult in HH	10.72	10.94	-0.23*	-0.23*	-0.23*
			(-1.86)	(-1.87)	(-1.92)
Number of productive adults in HH	3.52	3.49	0.022	0.022	0.020
			(0.34)	(0.34)	(0.31)
Land Size at Baseline	2.30	2.36	-0.077	-0.077	-0.076
			(-0.87)	(-0.87)	(-0.87)
2007 asset index (2016 expend. weighted) consumption expenditure data	13.07	12.96	0.11	0.11	0.077
			(0.52)	(0.51)	(0.36)
2007 asset index (PCA weighted)	1.54	1.49	0.056	0.055	0.048
			(1.38)	(1.39)	(1.21)
Estimated 2007 soil organic carbon (plot avg.)	24.38	22.97	1.42***	1.40***	1.36***
			(4.63)	(5.12)	(6.12)
Estimated 207 soil erosion (plot avg.)	0.36	0.36	0.00055	0.00044	0.00021
			(0.18)	(0.15)	(0.07)
Elevation (hh level)	1559.85	1529.87	29.6***	29.4***	-0.027
			(3.66)	(4.09)	(-0.61)
HH distance from tarmac road (km)	3.08	2.96	0.12*	0.12*	0.12*
			(1.67)	(1.68)	(1.77)
Observations	**1411**	**1386**	**2797**	**2797**	**2797**

z statistics in parenthesis; VSZ=Village Sampling Zone
* p<0.1, ** p<0.05, *** p<0.01

A2.3: Comparison of HHs in Vi Groups and Non-program Areas—Binary Characteristics

Characteristic	Vi-group Mean	Non-program Mean	Difference (raw)	Difference (net of county)	Difference (net of VSZ)
Respondent Female	0.60	0.60	0.00040	0.0038	0.033
			(0.020)	(0.07)	(0.65)
Respondent Christian	0.99	0.99	-0.0023	-0.14	-0.11
			(-0.72)	(-0.78)	(-0.56)
Respondent Luhya	0.96	0.98	-0.013*	-0.20**	-0.22**
			(-1.87)	(-1.99)	(-2.11)
Respondent Married	0.76	0.73	0.026	0.069	0.051
			(1.47)	(1.25)	(0.92)
Res. Married by Spouse Elsewhere	0.07	0.08	-0.0037	-0.017	-0.0048
			(-0.34)	(-0.23)	(-0.06)
Respondent Widowed	0.14	0.16	-0.016	-0.058	-0.045
			(-1.13)	(-0.92)	(-0.71)
Respondent is Head	0.51	0.54	-0.030	-0.071	-0.087*
			(-1.47)	(-1.40)	(-1.71)
Respondent is Spouse of Head	0.38	0.37	0.0087	0.022	0.045
			(0.44)	(0.42)	(0.86)
Respondent Good Health	0.97	0.98	-0.013**	-0.21**	-0.23**
			(-1.98)	(-1.98)	(-2.09)
Respondent Literate	0.89	0.89	0.0036	0.019	0.027
			(0.28)	(0.28)	(0.39)
Respondent in School	0.03	0.03	-0.00091	-0.013	-0.043
			(-0.14)	(-0.13)	(-0.40)
Respondent has Official Role	0.51	0.46	0.042**	0.11**	0.086*
			(2.06)	(2.09)	(1.68)
Respondent Farmer (main occupation)	0.87	0.86	0.019	0.088	0.094
			(1.38)	(1.38)	(1.46)
Respondent Primarily Business	0.05	0.06	-0.0067	-0.058	-0.063
			(-0.71)	(-0.71)	(-0.77)
Respondent Employed	0.05	0.06	-0.013	-0.12	-0.15*
			(-1.41)	(-1.38)	(-1.68)
Respondent has Tech. Skills	0.07	0.04	0.022**	0.20**	0.16**
			(2.40)	(2.35)	(1.96)
HH was farming in 2007	0.99	0.98	0.0027	0.062	0.062
			(0.53)	(0.48)	(0.48)
HH reared livestock in 2007	0.61	0.57	0.045**	0.12**	0.11**
			(2.26)	(2.36)	(2.15)
HH ran off-farm business in 2007	0.29	0.31	-0.020	-0.055	-0.055
			(-1.08)	(-1.02)	(-1.01)
HH member employed in 2007	0.16	0.13	0.030**	0.14**	0.15**
			(2.11)	(2.20)	(2.34)
HH used irrigation in 2007	0.06	0.04	0.013	0.12	0.100
			(1.44)	(1.39)	(1.14)
Head is productive	0.97	0.97	-0.0088	-0.12	-0.17*
			(-1.27)	(-1.22)	(-1.69)
all adults in HH over 59	0.06	0.05	0.0042	0.033	0.054
			(0.45)	(0.39)	(0.65)
Head is 60 or older	0.33	0.29	0.045**	0.13**	0.15***
			(2.42)	(2.46)	(2.80)
female headed HH	0.21	0.22	-0.0100	-0.025	-0.0097
			(-0.60)	(-0.43)	(-0.17)
Literate adult in HH	0.98	0.97	0.0082	0.15	0.14
			(1.33)	(1.33)	(1.26)
Female literate adult in HH	0.89	0.89	0.0013	0.000085	0.0091
			(0.10)	(0.00)	(0.13)
Head is literate	0.90	0.90	-0.0012	-0.0054	-0.0035
			(-0.096)	(-0.08)	(-0.05)
HH located on tarmac road	0.04	0.05	-0.010	-0.12	-0.11
			(-1.27)	(-1.35)	(-1.19)
HH had formal title to main parcel (07)	0.35	0.32	0.035*	0.11**	0.11**
			(1.85)	(2.00)	(2.14)
Respondent owned main parcel (07)	0.48	0.54	-0.061***	-0.15***	-0.17***
			(-2.99)	(-3.01)	(-3.37)
Observations	**1094**	**1386**	**2480**	**2480**	**2480**

z statistics in second column; VSZ=Village Sampling Zone

* p<0.1, ** p<0.05, *** p<0.01

Probit regression used for net of county and VSZ differences, so coefficients are not directly interpretable, only the *t*-statistics

A2.4: Comparison of HHs in Vi Groups and Non-program Areas—Continuous Characteristics

Characteristic	Program Mean	Non-program Mean	Difference (raw)	Difference (net of county)	Difference (net of VSZ)
Age of respondent	49.82	48.84	0.97*	1.01*	1.24**
			(1.79)	(1.88)	(2.28)
Years of education of respondent	8.59	8.61	-0.017	-0.0084	-0.027
			(-0.12)	(-0.06)	(-0.18)
Household size	6.64	6.65	-0.019	-0.030	-0.047
			(-0.19)	(-0.29)	(-0.46)
Number of children in HH	3.63	3.71	-0.083	-0.094	-0.10
			(-1.00)	(-1.14)	(-1.24)
Number of adults in HH	3.01	2.94	0.064	0.065	0.056
			(1.08)	(1.10)	(0.95)
Years of education head	9.07	9.20	-0.13	-0.12	-0.12
			(-0.88)	(-0.79)	(-0.80)
Highest years of educ. of any adult in HH	10.74	10.94	-0.21	-0.20	-0.21*
			(-1.63)	(-1.52)	(-1.65)
Number of productive adults in HH	3.58	3.49	0.075	0.076	0.066
			(1.09)	(1.11)	(0.96)
Land Size at Baseline	2.39	2.36	0.014	0.016	-0.015
			(0.14)	(0.17)	(-0.16)
2007 asset index (2016 expend. weighted) consumption expenditure data	13.10	12.96	0.14	0.19	0.16
			(0.62)	(0.85)	(0.72)
2007 asset index (PCA weighted)	1.54	1.49	0.049	0.061	0.059
			(1.15)	(1.44)	(1.39)
Distance to tarmac road (village level)	0.03	0.03	0.0012*	0.0011	0.0012*
			(1.66)	(1.58)	(1.84)
On road (village level)	0.02	0.03	-0.012*	-0.013*	-0.018***
			(-1.78)	(-1.89)	(-2.67)
Estimated 2007 soil organic carbon (plots)	23.10	22.97	0.14	0.41	0.89***
			(0.44)	(1.42)	(3.73)
Estimated 207 soil erosion prevalence (plots)	0.36	0.36	-0.0020	-0.00025	-0.00094
			(-0.60)	(-0.08)	(-0.29)
Estimated soil organic carbon (village level)	23.77	26.46	-2.70***	-2.52***	-1.99***
			(-10.6)	(-10.72)	(-10.17)
Elevation (hh level)	1528.96	1529.87	-1.45	5.27	19.1***
			(-0.17)	(0.67)	(2.72)
HH distance from tarmac road (km)	3.01	2.96	0.047	0.041	0.054
			(0.61)	(0.53)	(0.73)
Observations	**1094**	**1386**	**2480**	**2480**	**2480**

z statistics in parenthesis; VSZ=Village Sampling Zone
* p<0.1, ** p<0.05, *** p<0.01

Annex 3: Programme Exposure Tables

Comparison of HHs in Program and Non-program Areas—Group Participation

Characteristic	Program Mean	Non-program Mean	Difference (raw)	Difference (net of county)	Difference (net of VSZ)
>1 group 16	0.48	0.44	0.038**	0.096**	0.097**
			(2.03)	(2.03)	(2.04)
>1 group 07	0.10	0.08	0.028***	0.18***	0.18***
			(2.64)	(2.63)	(2.64)
Crop Production 16	0.77	0.71	0.060***	0.19***	0.20***
			(3.60)	(3.66)	(3.80)
Crop Production 07	0.62	0.55	0.067***	0.18***	0.19***
			(3.62)	(3.66)	(3.91)
Dairy 16	0.39	0.44	-0.055***	-0.14***	-0.15***
			(-2.97)	(-2.99)	(-3.12)
Dairy 07	0.21	0.24	-0.031**	-0.10**	-0.10**
			(-1.97)	(-1.98)	(-1.98)
Poultry 16	0.52	0.55	-0.034*	-0.086*	-0.092*
			(-1.79)	(-1.81)	(-1.94)
Poultry 07	0.29	0.31	-0.020	-0.056	-0.060
			(-1.13)	(-1.13)	(-1.21)
Other Livestock 16	0.11	0.11	0.0088	0.046	0.029
			(0.74)	(0.74)	(0.46)
Other Livestock 07	0.06	0.06	-0.0046	-0.040	-0.038
			(-0.52)	(-0.52)	(-0.50)
Marketing Crops or Livestock 16	0.13	0.14	-0.013	-0.060	-0.059
			(-1.01)	(-1.02)	(-1.00)
Marketing Crops or Livestock 07	0.07	0.09	-0.016	-0.11	-0.11
			(-1.59)	(-1.58)	(-1.56)
Soil & Water Conservation 16	0.14	0.04	0.094***	0.62***	0.61***
			(8.61)	(8.50)	(8.33)
Soil & Water Conservation 07	0.09	0.02	0.070***	0.72***	0.73***
			(8.12)	(7.83)	(7.72)
Tree Planting & Mgt. 16	0.54	0.33	0.21***	0.53***	0.54***
			(11.0)	(10.97)	(11.14)
Tree Planting & Mgt. 07	0.33	0.18	0.15***	0.47***	0.47***
			(8.97)	(8.91)	(8.92)
Water Use 16	0.02	0.02	0.0010	0.020	0.018
			(0.18)	(0.18)	(0.17)
Water Use 07	0.01	0.01	0.0020	0.091	0.084
			(0.59)	(0.59)	(0.54)
Microfinance or Savings 16	0.88	0.88	-0.0072	-0.034	-0.033
			(-0.58)	(-0.55)	(-0.53)
Microfinance or Savings 07	0.80	0.81	-0.019	-0.071	-0.066
			(-1.29)	(-1.30)	(-1.21)
Religious Activities 16	0.05	0.04	0.0078	0.083	0.080
			(0.99)	(1.00)	(0.97)
Religious Activities 07	0.02	0.03	-0.0062	-0.11	-0.11
			(-1.04)	(-1.04)	(-1.09)
Other 16	0.17	0.16	0.00063	-0.00047	-0.010
			(0.045)	(-0.01)	(-0.18)
Other 07	0.10	0.10	0.0040	0.026	0.025
			(0.36)	(0.40)	(0.37)
Observations	**1411**	**1386**	**2797**	**2797**	**2797**

z statistics in parenthesis; VSZ=Village Sampling Zone
* p<0.1, ** p<0.05, *** p<0.01
Probit regression used for net of county and VSZ differences, so coefficients are not directly interpretable, only the *t*-statistics

Comparison of HHs in Program and Non-program Areas—Training

Characteristic	Program Mean	Non-program Mean	Difference (raw)	Difference (net of county)	Difference (net of VSZ)
Significant training in at least 1 area 16	0.59	0.48	0.10***	0.26***	0.27***
			(5.52)	(5.52)	(5.57)
Significant training in at least 1 area 07	0.30	0.19	0.11***	0.34***	0.35***
			(6.54)	(6.52)	(6.60)
Significant training in at least 2 areas 16	0.36	0.24	0.11***	0.33***	0.33***
			(6.59)	(6.58)	(6.59)
Significant training in at least 2 areas 07	0.14	0.07	0.070***	0.40***	0.40***
			(6.06)	(6.00)	(6.00)
Crop training & impl. 16	0.37	0.34	0.036**	0.096**	0.098**
			(1.98)	(1.98)	(2.02)
Crop training & impl. 07	0.16	0.11	0.045***	0.20***	0.21***
			(3.44)	(3.42)	(3.55)
Dairy training & impl. 16	0.14	0.12	0.016	0.073	0.072
			(1.21)	(1.21)	(1.20)
Dairy training & impl. 07	0.05	0.03	0.018**	0.21**	0.20**
			(2.39)	(2.38)	(2.32)
Soil fertility training & impl. 16	0.13	0.06	0.070***	0.44***	0.44***
			(6.43)	(6.36)	(6.27)
Soil fertility training & impl. 07	0.05	0.02	0.036***	0.48***	0.50***
			(5.03)	(4.91)	(5.04)
Tree planting & mgt. training & impl. 16	0.27	0.12	0.16***	0.59***	0.60***
			(10.5)	(10.47)	(10.46)
Tree planting & mgt. training & impl. 07	0.12	0.03	0.084***	0.67***	0.66***
			(8.55)	(8.39)	(8.22)
Poultry mgt. training & impl. 16	0.16	0.12	0.036***	0.16***	0.16***
			(2.73)	(2.78)	(2.69)
Poultry mgt. training & impl. 07	0.05	0.05	0.0077	0.075	0.069
			(0.92)	(0.94)	(0.86)
Other livestock training & impl. 16	0.03	0.01	0.020***	0.41***	0.40***
			(3.58)	(3.54)	(3.49)
Other livestock training & impl. 07	0.01	0.01	0.0027	0.15	0.16
			(0.91)	(0.90)	(0.94)
Marketing training & impl. 16	0.03	0.02	0.0067	0.11	0.10
			(1.12)	(1.12)	(1.02)
Marketing training & impl. 07	0.01	0.01	0.0049	0.24	0.27
			(1.49)	(1.48)	(1.59)
Crop processing training & impl. 16	0.01	0.01	0.0013	0.072	0.089
			(0.43)	(0.43)	(0.52)
Crop processing training & impl. 07	0.01	0.00	0.0057***	0	0
			(2.81)	(.)	(.)
Financial mgt. or savings training & impl. 16	0.11	0.08	0.032***	0.20***	0.20***
			(2.94)	(2.95)	(2.97)
Financial mgt. or savings training & impl. 07	0.05	0.03	0.017**	0.20**	0.22**
			(2.28)	(2.29)	(2.45)
Off-farm business training & impl. 16	0.02	0.02	0.0053	0.099	0.10
			(0.96)	(0.93)	(0.96)
Off-farm business training & impl. 07	0.01	0.00	0.0049	0.27	0.28
			(1.57)	(1.56)	(1.61)
Observations	**1411**	**1386**	**2797**	**2797**	**2797**

z statistics in parenthesis; VSZ=Village Sampling Zone

* p<0.1, ** p<0.05, *** p<0.01

Probit regression used for net of county and VSZ differences, so coefficients are not directly interpretable, only the *t*-statistics

Annex 4: Agroforestry Adoption Index

A4.1: Comparison of HHs—Specific Agroforestry Adaption Indicators

Dimension	Binary Indicators		PA Mean	Vi Group Mean	Non-Vi in PA Mean	Non-PA Mean	PA vs. non-PA (dif.)	Vi vs. non-PA (dif.)	Vi vs. non-Vi in PA (dif.)
Practice Up-take	*2 or more tree products from main parcel*	2007	0.37	0.37	0.37	0.31	0.063***	0.063***	0.0020
							(3.50)	(3.31)	(0.07)
		2016	0.54	0.55	0.50	0.45	0.089***	0.10***	0.056*
							(4.73)	(5.05)	(1.75)
		Change	0.17	0.18	0.13	0.14	0.027	0.039**	0.053*
							(1.47)	(2.00)	(1.73)
	Tree-based NRM within food plots of main parcel	2007	0.04	0.05	0.02	0.04	0.0086	0.015*	0.028**
							(1.16)	(1.84)	(2.16)
		2016	0.14	0.16	0.05	0.06	0.080***	0.10***	0.11***
							(7.13)	(8.55)	(5.00)
		Change	0.09	0.11	0.03	0.02	0.071***	0.089***	0.081***
							(6.32)	(7.28)	(3.66)
	At least one complementary AF practice	2007	0.18	0.18	0.15	0.12	0.060***	0.069***	0.040
							(4.53)	(4.87)	(1.63)
		2016	0.25	0.27	0.20	0.18	0.078***	0.093***	0.067**
							(5.07)	(5.64)	(2.41)
		Change	0.08	0.08	0.06	0.06	0.018	0.024*	0.027
							(1.54)	(1.91)	(1.24)
Intensity of Practice	*50 trees on one plot or 20 on all pots*	2007	0.19	0.20	0.16	0.17	0.018	0.027*	0.039
							(1.25)	(1.72)	(1.57)
		2016	0.36	0.39	0.25	0.26	0.098***	0.13***	0.14***
							(5.66)	(6.93)	(4.50)
		Change	0.16	0.19	0.09	0.08	0.080***	0.10***	0.097***
							(4.73)	(5.58)	(3.29)
	Tree products from farm plots for sale	2007	0.17	0.17	0.17	0.13	0.035***	0.035**	-0.0012
							(2.59)	(2.42)	(-0.05)
		2016	0.32	0.34	0.24	0.22	0.10***	0.12***	0.096***
							(6.03)	(6.85)	(3.25)
		Change	0.15	0.17	0.07	0.08	0.065***	0.087***	0.097***
							(3.66)	(4.56)	(3.12)
	Intensity of complementary practices	2007	0.03	0.03	0.03	0.02	0.0089	0.0094	0.0022
							(1.59)	(1.58)	(0.21)
		2016	0.08	0.08	0.07	0.06	0.022**	0.025**	0.012
							(2.31)	(2.42)	(0.66)
		Change	0.05	0.06	0.05	0.04	0.013	0.016	0.0094
							(1.40)	(1.51)	(0.53)
	Products for sale from complementary practices	2007	0.05	0.05	0.04	0.03	0.021***	0.023***	0.0079
							(2.91)	(2.97)	(0.57)
		2016	0.13	0.14	0.10	0.07	0.055***	0.065***	0.041*
							(4.85)	(5.29)	(1.92)
		Change	0.08	0.09	0.05	0.05	0.034***	0.041***	0.033*
							(3.33)	(3.81)	(1.74)
Tree Species	*1 or more leguminous shrubs*	2007	0.15	0.16	0.12	0.08	0.069***	0.079***	0.045**
							(5.67)	(6.09)	(1.97)
		2016	0.44	0.48	0.28	0.20	0.23***	0.28***	0.20***
							(13.55)	(15.30)	(6.52)
		Change	0.28	0.32	0.16	0.12	0.16***	0.20***	0.16***
							(9.48)	(10.88)	(4.89)
	3 or more long-term exotic species	2007	0.08	0.09	0.08	0.07	0.017*	0.019*	0.0093
							(1.66)	(1.75)	(0.53)
		2016	0.19	0.20	0.15	0.12	0.071***	0.082***	0.049*
							(5.22)	(5.63)	(1.95)
		Change	0.11	0.12	0.08	0.05	0.054***	0.063***	0.039
							(4.04)	(4.37)	(1.59)
	3 or more native species	2007	0.13	0.12	0.14	0.09	0.032***	0.029**	-0.013
							(2.74)	(2.36)	(-0.62)
		2016	0.16	0.17	0.14	0.11	0.057***	0.065***	0.037
							(4.40)	(4.72)	(1.57)
		Change	0.04	0.05	0.00	0.01	0.025*	0.036***	0.050**
							(1.96)	(2.66)	(2.28)
		Observations	**1411**	**1094**	**317**	**1386**	**2797**	**2480**	**1411**

t statistics in parenthesis; PA = Programme Area

* p<0.1, ** p<0.05, *** p<0.01

Figure A4.1: HH Comparsion on Agroforesty Index: Program Area, Vi Group Only & Non-program Area by Weighted Indicator

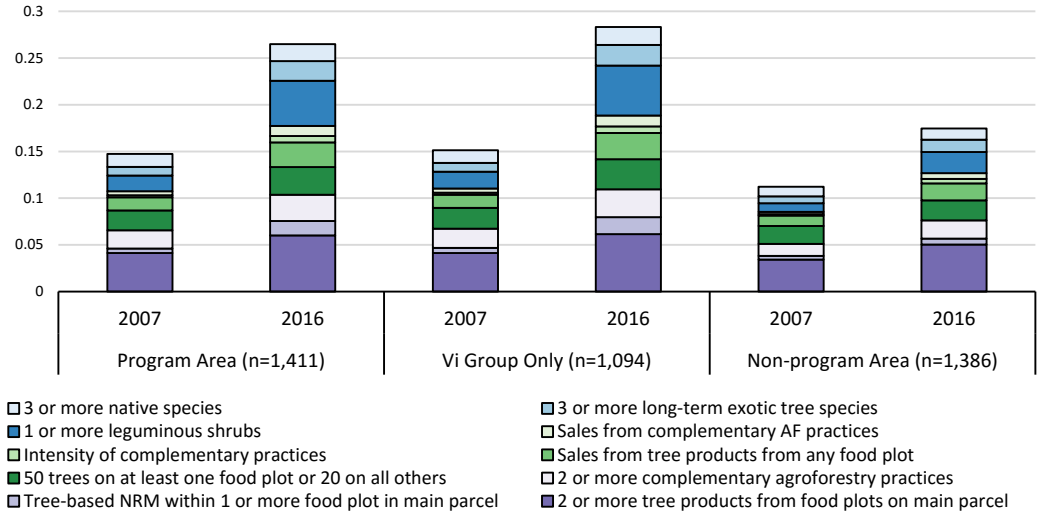
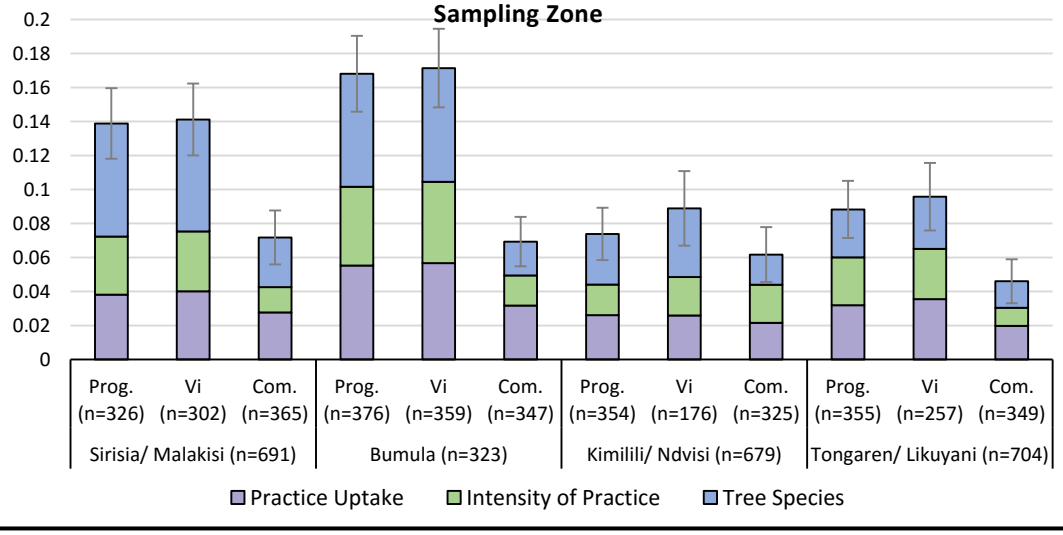


Figure A4.2: Disaggregated Comparison on Gains in Agroforesty Index: Program Area & Comparison Area by Weighted Dimension & Village Sampling Zone



A4.2: Comparison of HHs—Other Sustainable Land Management Adaption Indicators

Dimension	Binary Indicators		PA	Vi Group	Non-Vi in PA	Non-PA	PA vs. non-PA (dif.)	Vi vs. non-PA (dif.)	Vi vs. non-Vi in PA (dif.)
			Mean	Mean	Mean	Mean			
Crop Management	*Cover Crops*	2007	0.44	0.43	0.44	0.47	-0.034*	-0.036*	-0.0084
							(-1.80)	(-1.77)	(-0.26)
		2016	0.45	0.45	0.43	0.48	-0.036*	-0.031	0.023
							(-1.92)	(-1.54)	(0.71)
		Change	0.01	0.02	-0.01	0.01	-0.0024	0.0046	0.031
							(-0.14)	(0.25)	(1.06)
	Crop Rotation/ Sequential Cropping	2007	0.53	0.53	0.54	0.55	-0.019	-0.020	-0.0061
							(-1.01)	(-1.01)	(-0.19)
		2016	0.58	0.58	0.59	0.60	-0.018	-0.021	-0.0095
							(-0.99)	(-1.03)	(-0.30)
		Change	0.05	0.05	0.05	0.05	0.00052	-0.00023	-0.0034
							(0.03)	(-0.01)	(-0.12)
	Intercropping	2007	0.91	0.91	0.90	0.92	-0.0078	-0.0055	0.010
							(-0.73)	(-0.49)	(0.55)
		2016	0.93	0.93	0.92	0.93	-0.0052	-0.0040	0.0053
							(-0.54)	(-0.39)	(0.32)
		Change	0.02	0.02	0.02	0.02	0.0026	0.0015	-0.0047
							(0.22)	(0.12)	(-0.24)
	Relay cropping	2007	0.60	0.59	0.65	0.63	-0.029	-0.044**	-0.065**
							(-1.57)	(-2.21)	(-2.09)
		2016	0.66	0.66	0.66	0.67	-0.013	-0.013	-0.00026
							(-0.75)	(-0.70)	(-0.01)
		Change	0.06	0.07	0.01	0.04	0.016	0.030*	0.065**
							(0.93)	(1.68)	(2.21)
	Strip cropping	2007	0.04	0.04	0.05	0.05	-0.0073	-0.0087	-0.0062
							(-0.92)	(-1.02)	(-0.48)
		2016	0.08	0.09	0.05	0.09	-0.0051	0.0026	0.034**
							(-0.49)	(0.23)	(1.97)
		Change	0.04	0.05	0.01	0.04	0.0022	0.011	0.040**
							(0.24)	(1.13)	(2.52)
Soil & Water Conservation	*Mulching*	2007	0.26	0.26	0.26	0.23	0.030*	0.030*	0.0014
							(1.83)	(1.73)	(0.05)
		2016	0.31	0.31	0.30	0.28	0.031*	0.033*	0.0070
							(1.82)	(1.79)	(0.24)
		Change	0.05	0.05	0.04	0.04	0.0013	0.0026	0.0056
							(0.09)	(0.16)	(0.22)
	Reduced tillage	2007	0.00	0.00	0.00	0.00	0.00068	0.0013	0.0027
							(0.43)	(0.72)	(0.93)
		2016	0.01	0.02	0.01	0.02	-0.0024	-0.00033	0.0092
							(-0.53)	(-0.07)	(1.26)
		Change	0.01	0.01	0.01	0.01	-0.0031	-0.0016	0.0065
							(-0.69)	(-0.33)	(0.96)
	Terracing	2007	0.06	0.07	0.03	0.04	0.021**	0.031***	0.045***
							(2.49)	(3.39)	(2.98)
		2016	0.09	0.11	0.02	0.06	0.028***	0.047***	0.087***
							(2.80)	(4.28)	(4.80)
		Change	0.03	0.04	-0.00	0.02	0.0074	0.017*	0.042***
							(0.86)	(1.76)	(2.68)
	Trenches	2007	0.04	0.05	0.02	0.03	0.013*	0.020**	0.030**
							(1.85)	(2.55)	(2.31)
		2016	0.07	0.08	0.04	0.05	0.025***	0.034***	0.043**
							(2.70)	(3.45)	(2.58)
		Change	0.03	0.04	0.02	0.02	0.012	0.015	0.014
							(1.38)	(1.58)	(0.89)
	Contour Bunds	2007	0.06	0.06	0.06	0.05	0.0083	0.0089	0.0026
							(0.97)	(0.97)	(0.18)
		2016	0.07	0.07	0.05	0.08	-0.011	-0.0050	0.025
							(-1.08)	(-0.47)	(1.57)
		Change	0.01	0.01	-0.01	0.03	-0.019**	-0.014	0.022*
							(-2.40)	(-1.62)	(1.71)
Soil Fertility Management	*Fallow (not improved)*	2007	0.06	0.07	0.02	0.03	0.023***	0.036***	0.055***
							(2.95)	(4.05)	(3.67)
		2016	0.06	0.06	0.04	0.03	0.030***	0.035***	0.024*
							(3.97)	(4.37)	(1.65)
		Change	-0.00	-0.01	0.02	-0.01	0.0065	-0.00029	-0.030**
							(0.90)	(-0.04)	(-2.31)
	Fallow (improved with grass)	2007	0.02	0.02	0.00	0.01	0.011***	0.015***	0.019**
							(2.71)	(3.34)	(2.23)
		2016	0.02	0.03	0.00	0.02	0.0068	0.013**	0.026***
							(1.28)	(2.13)	(2.71)
		Change	0.01	0.01	0.00	0.01	-0.0044	-0.0028	0.0073
							(-0.83)	(-0.47)	(0.72)
	Animal manure	2007	0.51	0.50	0.54	0.50	0.0060	-0.0029	-0.039
							(0.32)	(-0.14)	(-1.24)
		2016	0.60	0.60	0.60	0.56	0.042**	0.043**	0.0071
							(2.24)	(2.18)	(0.23)
		Change	0.09	0.10	0.06	0.06	0.036**	0.046**	0.047
							(2.09)	(2.50)	(1.60)
	Composting	2007	0.10	0.10	0.10	0.08	0.023**	0.021*	-0.0063

						(2.12)	(1.88)	(-0.33)
	2016	0.18	0.18	0.16	0.14	0.038***	0.044***	0.026
						(2.74)	(2.96)	(1.07)
	Change	0.08	0.09	0.05	0.06	0.015	0.022*	0.032
						(1.22)	(1.69)	(1.43)
Green Manure	2007	0.20	0.19	0.21	0.17	0.031**	0.027*	-0.018
						(2.13)	(1.75)	(-0.69)
	2016	0.24	0.25	0.20	0.19	0.048***	0.060***	0.053*
						(3.08)	(3.59)	(1.96)
	Change	0.04	0.05	-0.02	0.02	0.017	0.032**	0.071***
						(1.23)	(2.27)	(2.94)
	Observations	**1411**	**1094**	**317**	**1386**	**2797**	**2480**	**1411**

t statistics in parenthesis; PA = Programme Area
* p<0.1, ** p<0.05, *** p<0.01

Annex 5: Differential Effect Estimates

A5.1 Program and Non-Program Area Comparison on AF Index and Dimensions, Differential Effects

	Overall Index	D1: Practice Uptake	D2: Intensity of Practice	D3: Tree Species
Overall Double Difference	0.055***	0.039***	0.047***	0.081***
	(0.0063)	(0.0087)	(0.0083)	(0.0090)
Observations	**2797**	**2797**	**2797**	**2797**
By Village Sampling Zone				
Sirisia/Malakisi*Program Area	0.067***	0.031	0.058***	0.11***
	(0.013)	(0.019)	(0.017)	(0.020)
Bumula*Program Area	0.099***	0.071***	0.086***	0.14***
	(0.014)	(0.019)	(0.018)	(0.018)
Kimilili/Ndivisi*Program Area	0.012	0.014	-0.013	0.036**
	(0.011)	(0.016)	(0.016)	(0.018)
Tongaren/Likuyani*Program Area	0.042***	0.037**	0.053***	0.038**
	(0.011)	(0.015)	(0.015)	(0.015)
Wald tests (F)				
Sirisia/Malakisi vs. Bumula	2.772*	2.165	1.240	1.083
p-value	0.096	0.141	0.266	0.298
Sirisia/Malakisi vs. Kimilili/Ndivisi	9.819***	0.518	9.282***	7.980***
p-value	0.002	0.472	0.002	0.005
Sirisia/Malakisi vs. Tongaren/Likuyani	2.081	0.045	0.057	8.591***
p-value	0.149	0.832	0.811	0.003
Bumula vs. Kimilili/Ndivisi	23.850***	5.454*	17.208***	16.938***
p-value	0.000	0.020	0.000	0.000
Bumula vs. Tongaren/Likuyani	10.544***	2.039	2.003	18.739***
p-value	0.001	0.153	0.157	0.000
Kimilili/Ndivisi vs. Tongaren/Likuyani	3.675*	1.115	9.054***	0.005
p-value	0.055	0.291	0.003	0.946
By Gender				
Female*Program Area	0.057***	0.045***	0.051***	0.075***
	(0.0072)	(0.0100)	(0.0095)	(0.011)
Male*Program Area	0.056***	0.032**	0.043***	0.091***
	(0.011)	(0.016)	(0.015)	(0.016)
Wald tests (F)				
Female vs. Male	0.010	0.441	0.175	0.701
p-value	0.920	0.506	0.675	0.402
By Sex of Head				
Female Headed*Program Area	0.067***	0.056***	0.070***	0.074***
	(0.012)	(0.016)	(0.015)	(0.018)
Male Headed*Program Area	0.052***	0.033***	0.040***	0.082***
	(0.0073)	(0.010)	(0.0097)	(0.010)
Wald tests (F)				
Female Headed vs. Male Headed	1.196	1.405	2.825*	0.149
p-value	0.274	0.236	0.093	0.699

	Overall Index	D1: Practice Uptake	D2: Intensity of Practice	D3: Tree Species
By Respondent Education				
Over 7 years*Program Area	0.055*** (0.0080)	0.035*** (0.011)	0.044*** (0.010)	0.086*** (0.011)
Under 8 years*Program Area	0.056*** (0.010)	0.045*** (0.014)	0.051*** (0.014)	0.071*** (0.015)
Wald tests (F)				
Under 8s vs. Over 7s *p*-value	0.001 0.973	0.284 0.594	0.153 0.696	0.643 0.423
By Baseline Wealth Status (Asset-based)				
Asset Poor*Program Area	0.064*** (0.0089)	0.052*** (0.012)	0.051*** (0.011)	0.090*** (0.012)
Asset Rich*Program Area	0.046*** (0.0088)	0.026** (0.012)	0.042*** (0.012)	0.071*** (0.013)
Wald tests (F)				
Asset Rich vs. Asset Poor *p*-value	2.079 0.149	2.278 0.131	0.325 0.568	1.052 0.305
By Landholding Size				
< 2 acre*Program Area	0.056*** (0.0083)	0.044*** (0.012)	0.052*** (0.010)	0.073*** (0.012)
≥ 2 acre*Program Area	0.054*** (0.0095)	0.032** (0.013)	0.040*** (0.013)	0.089*** (0.014)
Wald tests (F)				
Under 2 vs. 2 or overs *p*-value	0.053 0.819	0.431 0.511	0.544 0.461	0.705 0.401
By Dairy Producer				
Dairy*Program Area	0.039*** (0.0096)	0.016 (0.013)	0.037*** (0.013)	0.063*** (0.014)
No Dairy*Program Area	0.068*** (0.0083)	0.056*** (0.011)	0.054*** (0.011)	0.095*** (0.012)
Wald tests (F)				
Dairy vs. No dairy *p*-value	5.384** 0.020	5.139** 0.023	0.959 0.327	3.001* 0.083
By Official Position				
Official Position*Program Area	0.056*** (0.0096)	0.040*** (0.013)	0.034*** (0.013)	0.094*** (0.014)
No Official Position*Program Area	0.053*** (0.0081)	0.037*** (0.011)	0.056*** (0.011)	0.067*** (0.012)
Wald tests (F)				
Off. Pos. vs. No Off. Pos. *p*-value	0.053 0.818	0.033 0.857	1.735 0.188	2.312 0.128
By Land Owned by Respondent (formal or informal)				
Owned*Program Area	0.058***	0.033**	0.046***	0.094***

	Overall Index	D1: Practice Uptake	D2: Intensity of Practice	D3: Tree Species
	(0.011)	(0.015)	(0.014)	(0.015)
Not Owned*Program Area	0.057***	0.048***	0.050***	0.072***
	(0.0089)	(0.012)	(0.012)	(0.013)
Wald tests (F)				
Owned vs. Not Owned	0.069	0.413	0.019	1.820
p-value	0.793	0.521	0.889	0.178
By HH Land Title (2007)				
Title*Program Area	0.053***	0.046***	0.029**	0.083***
	(0.0098)	(0.013)	(0.013)	(0.013)
No Title*Program Area	0.060***	0.038***	0.058***	0.083***
	(0.0085)	(0.011)	(0.011)	(0.012)
Wald tests (F)				
Title vs. No Title	1.046	0.070	6.200	0.052
p-value	0.307	0.791	0.013	0.819

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with programme area (p=<0.1) used in all models; VSZ dummies used for fixed effects in all linear models

A5.2: Remote Sensing Derived Tree Cover & Veg. Cover Estimates, Differential Effects

	Est. % of tree cover across plots				***Est. % of veg. cover across plots***			
	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.
Overall OLS ITT estimates	4.99***	2.97***	5.11***	3.10***	-0.012	0.66	-0.097	0.56
	(0.61)	(0.78)	(0.63)	(0.81)	(0.56)	(0.58)	(0.56)	(0.59)
Observations	**2784**	**2784**	**2753**	**2753**	**2790**	**2790**	**2759**	**2759**
By Village Sampling Zone								
Sirisia/Malakisi*Program Area	6.43***	2.83**	6.16***	2.88**	-3.71**	-2.97**	-3.76**	-3.02**
	(0.86)	(1.31)	(0.89)	(1.42)	(1.51)	(1.49)	(1.52)	(1.50)
Bumula*Program Area	5.43***	3.07**	5.48***	2.68**	2.77***	4.01***	2.54**	3.79***
	(1.15)	(1.30)	(1.17)	(1.30)	(1.06)	(1.13)	(1.07)	(1.13)
Kimilili/Ndivisi*Program Area	1.73*	-1.51	1.90*	-0.89	0.31	0.85	0.20	0.70
	(1.03)	(1.55)	(1.03)	(1.57)	(1.00)	(0.96)	(1.02)	(0.99)
Tongaren/Likuyani*Program Area	6.00***	6.96***	6.52***	7.28***	0.54	0.63	0.61	0.67
	(1.59)	(1.97)	(1.63)	(2.06)	(0.62)	(0.72)	(0.64)	(0.74)
Wald tests (F)								
Sirisia/Malakisi vs. Bumula	0.483	0.017	0.214	0.010	12.209	13.786	11.414	13.130
p-value	0.487	0.897	0.644	0.919	0.001	0.000	0.001	0.000
Sirisia/Malakisi vs. Kimilili/Ndivisi	11.977	4.426	9.556	3.037	4.987	4.805	4.732	4.496
p-value	0.001	0.036	0.002	0.082	0.026	0.029	0.030	0.035
Sirisia/Malakisi vs. Tongaren/Likuyani	0.059	3.079	0.037	3.130	6.802	4.716	7.015	4.904
p-value	0.809	0.080	0.848	0.078	0.009	0.030	0.008	0.027
Bumula vs. Kimilili/Ndivisi	5.865	5.130	5.371	3.040	2.767	4.631	2.451	4.326
p-value	0.016	0.024	0.021	0.082	0.097	0.032	0.118	0.038
Bumula vs. Tongaren/Likuyani	0.083	2.779	0.272	3.664	3.285	6.427	2.397	5.384
p-value	0.774	0.096	0.602	0.056	0.071	0.012	0.122	0.021
Kimilili/Ndivisi vs. Tongaren/Likuyani	5.087	10.917	5.732	9.598	0.038	0.033	0.113	0.001
p-value	0.025	0.001	0.017	0.002	0.845	0.856	0.737	0.982
By Gender								
Female*Program Area	4.39***	2.01**	4.60***	2.01**	-0.44	0.23	-0.44	0.21
	(0.70)	(0.90)	(0.71)	(0.94)	(0.59)	(0.61)	(0.60)	(0.62)
Male*Program Area	5.93***	4.48***	5.91***	4.82***	0.66	1.34	0.45	1.11
	(0.80)	(1.15)	(0.83)	(1.22)	(0.79)	(0.81)	(0.81)	(0.82)
Wald tests (F)								
Female vs. Male	3.296	3.606	2.218	4.108	1.946	1.993	1.243	1.286
p-value	0.070	0.058	0.137	0.043	0.164	0.159	0.266	0.257
By Sex of Head								
Female Headed*Program Area	3.61***	0.67	3.67***	0.57	-0.097	0.70	-0.12	0.66
	(0.90)	(1.39)	(0.91)	(1.42)	(0.74)	(0.76)	(0.76)	(0.78)
Male Headed*Program Area	5.38***	3.60***	5.51***	3.80***	0.020	0.65	-0.080	0.54
	(0.68)	(0.88)	(0.69)	(0.92)	(0.60)	(0.63)	(0.60)	(0.63)
Wald tests (F)								
Female Headed vs. Male Headed	3.358	3.433	3.417	3.981	0.027	0.004	0.003	0.024
p-value	0.068	0.065	0.065	0.047	0.870	0.950	0.953	0.876
By Respondent Education								
Over 7 years*Program Area	5.07***	3.37***	5.21***	3.56***	0.19	0.82	0.028	0.65
	(0.67)	(0.89)	(0.68)	(0.94)	(0.56)	(0.58)	(0.57)	(0.59)

	Est. % of tree cover across plots				Est. % of veg. cover across plots			
	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.
Under 8 years*Program Area	4.84***	2.21**	4.91***	2.23*	-0.38	0.37	-0.33	0.39
	(0.85)	(1.11)	(0.87)	(1.18)	(0.73)	(0.76)	(0.73)	(0.77)
Wald tests (F)								
Under 8s vs. Over 7s	0.070	0.906	0.108	1.016	0.966	0.577	0.368	0.187
p-value	0.791	0.342	0.743	0.314	0.326	0.448	0.544	0.666
By Baseline Wealth Status (Asset-based)								
Asset Poor*Program Area	4.44***	1.65*	4.52***	1.59	0.27	1.01	0.20	0.93
	(0.70)	(0.94)	(0.72)	(0.99)	(0.74)	(0.77)	(0.75)	(0.78)
Asset Rich*Program Area	5.53***	4.24***	5.69***	4.56***	-0.32	0.28	-0.41	0.17
	(0.76)	(1.03)	(0.78)	(1.08)	(0.56)	(0.59)	(0.57)	(0.60)
Wald tests (F)								
Asset Rich vs. Asset Poor	1.866	4.600	2.007	5.191	0.720	1.050	0.727	1.090
p-value	0.173	0.033	0.157	0.023	0.397	0.306	0.394	0.297
By Landholding Size								
< 2 acre*Program Area	4.43***	2.93***	4.66***	3.11***	0.48	1.19*	0.37	1.08*
	(0.73)	(1.01)	(0.75)	(1.01)	(0.61)	(0.64)	(0.61)	(0.65)
≥ 2 acre*Program Area	5.67***	2.99***	5.66***	3.10***	-0.53	0.092	-0.57	0.013
	(0.71)	(0.98)	(0.73)	(1.06)	(0.67)	(0.70)	(0.68)	(0.70)
Wald tests (F)								
Under 2 vs. 2 or overs	2.660	0.002	1.595	0.000	2.683	2.885	2.263	2.687
p-value	0.104	0.962	0.207	0.996	0.102	0.090	0.133	0.102
By Dairy Producer								
Dairy*Program Area	5.43***	3.29***	5.55***	3.27***	-0.24	0.32	-0.26	0.30
	(0.74)	(1.06)	(0.76)	(1.11)	(0.64)	(0.66)	(0.65)	(0.67)
No Dairy*Program Area	4.66***	2.72***	4.76***	2.97***	0.17	0.92	0.034	0.77
	(0.71)	(0.92)	(0.73)	(0.95)	(0.65)	(0.68)	(0.65)	(0.68)
Wald tests (F)								
Dairy vs. No dairy	0.977	0.216	0.970	0.056	0.402	0.785	0.198	0.480
p-value	0.323	0.642	0.325	0.813	0.526	0.376	0.657	0.489
By Official Position								
Official Position*Program Area	4.67***	2.93***	4.87***	3.38***	-0.40	0.36	-0.46	0.30
	(0.73)	(1.02)	(0.76)	(1.07)	(0.61)	(0.66)	(0.62)	(0.66)
No Official Position*Program Area	5.30***	3.01***	5.34***	2.87***	0.33	0.92	0.22	0.78
	(0.72)	(0.92)	(0.72)	(0.96)	(0.65)	(0.66)	(0.67)	(0.68)
Wald tests (F)								
Off. Pos. vs. No Off. Pos.	0.666	0.004	0.347	0.183	1.442	0.845	1.126	0.547
p-value	0.415	0.947	0.556	0.669	0.231	0.359	0.289	0.460
By Land Owned by Respondent (formal or informal)								
Owned*Program Area	4.87***	3.21***	5.02***	3.62***	-0.031	0.73	-0.22	0.51
	(0.74)	(1.03)	(0.77)	(1.08)	(0.68)	(0.72)	(0.69)	(0.73)
Not Owned*Program Area	5.12***	2.73***	5.20***	2.58**	0.0068	0.59	0.024	0.60
	(0.76)	(1.00)	(0.78)	(1.04)	(0.62)	(0.64)	(0.63)	(0.64)
Wald tests (F)								
Owned vs. Not Owned	0.082	0.137	0.043	0.593	0.003	0.043	0.128	0.014
p-value	0.775	0.712	0.836	0.442	0.955	0.836	0.721	0.905

	Est. % of tree cover across plots				***Est. % of veg. cover across plots***			
	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.
By HH Land Title (2007)								
Title*Program Area	4.74***	3.47***	4.83***	3.35***	-0.76	-0.26	-0.77	-0.30
	(0.83)	(1.16)	(0.86)	(1.25)	(0.64)	(0.67)	(0.66)	(0.68)
No Title*Program Area	5.10***	2.71***	5.23***	2.97***	0.36	1.11*	0.24	0.98
	(0.67)	(0.88)	(0.68)	(0.91)	(0.63)	(0.66)	(0.64)	(0.66)
Wald tests (F)								
Title vs. No Title	0.190	0.360	0.213	0.078	3.003	4.308	2.380	3.621
p-value	0.663	0.549	0.644	0.780	0.084	0.039	0.124	0.058

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with programme area (p=<0.1) used in all models; VSZ dummies used for fixed effects in all linear models

A5.3: Remote Sensing Derived Land Health Estimates for Main Parcel Plots, Differential Effects

	Grams of organic carbon/kg of soil				*% of plot area with significant soil erosion*			
	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.
Overall OLS ITT estimates	1.56***	1.15***	1.57***	1.15***	0.016*	0.015	0.017**	0.015
	(0.46)	(0.43)	(0.48)	(0.44)	(0.0084)	(0.0096)	(0.0085)	(0.0097)
Observations	**2790**	**2790**	**2759**	**2759**	**2790**	**2790**	**2759**	**2759**
By Village Sampling Zone								
Sirisia/Malakisi*Program Area	-1.13	-1.15	-1.28	-1.28	0.032*	0.037*	0.035**	0.038*
	(1.08)	(0.87)	(1.09)	(0.87)	(0.018)	(0.020)	(0.018)	(0.020)
Bumula*Program Area	0.68	1.04*	0.76	1.17**	0.0065	0.016	0.0091	0.017
	(0.51)	(0.59)	(0.52)	(0.59)	(0.015)	(0.017)	(0.016)	(0.018)
Kimilili/Ndivisi*Program Area	4.77***	2.24***	4.80***	2.13***	-0.053***	-0.061***	-0.055***	-0.062***
	(0.78)	(0.67)	(0.80)	(0.69)	(0.019)	(0.022)	(0.019)	(0.022)
Tongaren/Likuyani*Program Area	2.24**	2.47**	2.38**	2.59**	0.074***	0.065***	0.076***	0.064***
	(1.08)	(1.15)	(1.12)	(1.20)	(0.013)	(0.015)	(0.013)	(0.015)
Wald tests (F)								
Sirisia/Malakisi vs. Bumula	2.276	4.301	2.820	5.341	1.197	0.643	1.198	0.635
p-value	0.132	0.039	0.094	0.021	0.275	0.423	0.274	0.426
Sirisia/Malakisi vs. Kimilili/Ndivisi	18.486	9.563	19.092	9.483	10.810	10.999	11.922	11.016
p-value	0.000	0.002	0.000	0.002	0.001	0.001	0.001	0.001
Sirisia/Malakisi vs. Tongaren/Likuyani	4.846	6.325	5.483	6.838	3.703	1.282	3.436	1.082
p-value	0.028	0.012	0.020	0.009	0.055	0.258	0.064	0.299
Bumula vs. Kimilili/Ndivisi	18.912	1.845	17.692	1.124	5.992	7.634	6.757	7.592
p-value	0.000	0.175	0.000	0.290	0.015	0.006	0.010	0.006
Bumula vs. Tongaren/Likuyani	1.687	1.240	1.716	1.122	11.322	4.677	10.783	4.288
p-value	0.195	0.266	0.191	0.290	0.001	0.031	0.001	0.039
Kimilili/Ndivisi vs. Tongaren/Likuyani	3.701	0.031	3.238	0.111	30.292	22.668	31.686	22.376
p-value	0.055	0.859	0.073	0.740	0.000	0.000	0.000	0.000
By Gender								
Female*Program Area	1.29**	0.83*	1.32**	0.80	0.015	0.011	0.015	0.0099
	(0.51)	(0.48)	(0.53)	(0.49)	(0.0092)	(0.011)	(0.0094)	(0.011)
Male*Program Area	1.97***	1.64***	2.02***	1.71***	0.016	0.021	0.018	0.024*
	(0.57)	(0.54)	(0.58)	(0.57)	(0.012)	(0.013)	(0.012)	(0.013)
Wald tests (F)								
Female vs. Male	1.513	2.243	1.448	2.589	0.011	0.603	0.084	0.942
p-value	0.219	0.135	0.229	0.108	0.915	0.438	0.772	0.332
By Sex of Head								
Female Headed*Program Area	1.30**	0.97	1.23**	0.86	0.030***	0.026*	0.026**	0.023*
	(0.60)	(0.60)	(0.62)	(0.61)	(0.011)	(0.014)	(0.011)	(0.014)
Male Headed*Program Area	1.63***	1.20**	1.71***	1.24**	0.011	0.012	0.013	0.013
	(0.50)	(0.47)	(0.51)	(0.49)	(0.0092)	(0.010)	(0.0094)	(0.010)
Wald tests (F)								
Female Headed vs. Male Headed	0.297	0.137	0.567	0.343	2.770	1.087	1.110	0.564
p-value	0.586	0.711	0.452	0.558	0.097	0.298	0.293	0.453
By Respondent Education								
Over 7 years*Program Area	1.50***	1.00**	1.54***	1.00**	0.012	0.013	0.013	0.013
	(0.52)	(0.48)	(0.52)	(0.50)	(0.0091)	(0.011)	(0.0094)	(0.011)
Under 8 years*Program Area	1.62***	1.37***	1.73***	1.37**	0.022**	0.020	0.022**	0.020*
	(0.53)	(0.52)	(0.55)	(0.53)	(0.011)	(0.012)	(0.011)	(0.012)

	Grams of organic carbon/kg of soil				***% of plot area with significant soil erosion***			
	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.	2016 plot avg.	Dif. plot avg.	2016 food plot avg.	Dif. food plot avg.
Wald tests (F)								
Under 8s vs. Over 7s	0.052	0.534	0.138	0.478	1.026	0.363	0.949	0.391
p-value	0.820	0.465	0.711	0.490	0.312	0.547	0.331	0.532
By Baseline Wealth Status (Asset-based)								
Asset Poor*Program Area	0.75	0.42	0.75	0.38	0.019*	0.021*	0.019*	0.020*
	(0.56)	(0.51)	(0.57)	(0.51)	(0.010)	(0.011)	(0.011)	(0.012)
Asset Rich*Program Area	2.36***	1.87***	2.45***	1.92***	0.012	0.0091	0.014	0.011
	(0.53)	(0.52)	(0.54)	(0.54)	(0.0094)	(0.011)	(0.0095)	(0.011)
Wald tests (F)								
Asset Rich vs. Asset Poor	7.834	6.540	8.135	6.972	0.555	0.994	0.227	0.597
p-value	0.005	0.011	0.005	0.009	0.457	0.319	0.634	0.440
By Landholding Size								
< 2 acre*Program Area	1.60***	1.08**	1.54***	0.98*	0.018*	0.018	0.017*	0.018
	(0.54)	(0.50)	(0.55)	(0.51)	(0.0096)	(0.011)	(0.0099)	(0.011)
≥ 2 acre*Program Area	1.51***	1.22**	1.66***	1.34**	0.013	0.011	0.015	0.012
	(0.53)	(0.52)	(0.54)	(0.54)	(0.010)	(0.011)	(0.010)	(0.011)
Wald tests (F)								
Under 2 vs. 2 or overs	0.033	0.062	0.050	0.405	0.324	0.303	0.039	0.245
p-value	0.857	0.804	0.824	0.525	0.570	0.583	0.843	0.621
By Dairy Producer								
Dairy*Program Area	1.73***	1.24**	1.89***	1.37**	0.012	0.014	0.013	0.015
	(0.57)	(0.57)	(0.59)	(0.59)	(0.010)	(0.012)	(0.011)	(0.012)
No Dairy*Program Area	1.42***	1.07**	1.37***	0.97**	0.018**	0.015	0.019*	0.016
	(0.50)	(0.47)	(0.51)	(0.47)	(0.0093)	(0.010)	(0.0095)	(0.010)
Wald tests (F)								
Dairy vs. No dairy	0.320	0.081	0.838	0.473	0.386	0.011	0.312	0.003
p-value	0.572	0.776	0.361	0.492	0.535	0.917	0.577	0.956
By Official Position								
Official Position*Program Area	1.56***	1.13**	1.61***	1.15**	0.0086	0.0039	0.0083	0.0034
	(0.52)	(0.49)	(0.53)	(0.51)	(0.0094)	(0.011)	(0.0096)	(0.011)
No Official Position*Program Area	1.59***	1.18**	1.62***	1.17**	0.021**	0.024**	0.023**	0.025**
	(0.52)	(0.49)	(0.53)	(0.51)	(0.0096)	(0.011)	(0.0097)	(0.011)
Wald tests (F)								
Off. Pos. vs. No Off. Pos.	0.006	0.012	0.000	0.001	2.180	3.779	2.574	4.216
p-value	0.939	0.912	0.992	0.976	0.141	0.053	0.109	0.041

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with programme area (p=<0.1) used in all models; VSZ dummies used for fixed effects in all linear models

A5.5: Tree Product Sales, Differential Effects

	Sale of AF Products Ksh (2016)	Sale of AF over Ksh 1000 (2016)	Sale of AF Products (log) (2016)	Sale of AF Products Ksh (difference)	Sale of AF over Ksh 1000 (difference)	Sale of AF Products (log) (difference)
Overall OLS Estimates	3428.0***	0.34***	1.06***	3384.9***	0.36***	0.043***
	(886.8)	(0.061)	(0.18)	(858.2)	(0.062)	(0.014)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
By Village Sampling Zone						
Sirisia/Malakisi*Program Area	4418.4**	0.44***	1.37***	4383.8**	0.46***	0.033**
	(1925.7)	(0.11)	(0.34)	(1926.1)	(0.11)	(0.017)
Bumula*Program Area	6305.3***	0.50***	1.61***	6237.7***	0.58***	0.080**
	(1539.6)	(0.12)	(0.36)	(1501.9)	(0.12)	(0.034)
Kimilili/Ndivisi*Program Area	-508.3	0.054	0.22	-547.7	0.019	-0.0086
	(1289.6)	(0.14)	(0.37)	(1272.1)	(0.14)	(0.012)
Tongaren/Likuyani*Program Area	3000.6	0.32***	0.94***	2968.6	0.32***	0.062*
	(1859.9)	(0.11)	(0.33)	(1831.3)	(0.12)	(0.037)
Wald tests (F)						
Sirisia/Malakisi vs. Bumula	0.576	0.120	0.231	0.563	0.550	1.413
p-value	0.448	0.726	0.631	0.454	0.457	0.235
Sirisia/Malakisi vs. Kimilili/Ndivisi	5.208	4.870	5.584	5.230	6.010	4.455
p-value	0.023	0.027	0.019	0.023	0.014	0.035
Sirisia/Malakisi vs. Tongaren/Likuyani	0.285	0.610	0.862	0.286	0.670	0.616
p-value	0.594	0.434	0.354	0.593	0.414	0.433
Bumula vs. Kimilili/Ndivisi	11.455	5.950	7.233	11.412	9.310	5.371
p-value	0.001	0.015	0.007	0.001	0.002	0.021
Bumula vs. Tongaren/Likuyani	1.955	1.200	1.939	1.933	2.280	0.124
p-value	0.163	0.273	0.164	0.165	0.131	0.724
Kimilili/Ndivisi vs. Tongaren/Likuyani	2.382	2.140	2.099	2.411	2.840	3.233
p-value	0.123	0.143	0.148	0.121	0.092	0.073
By Gender						
Female*Program Area	2838.3***	0.39***	1.02***	2836.9***	0.39***	0.048**
	(711.0)	(0.083)	(0.20)	(717.5)	(0.085)	(0.021)
Male*Program Area	4359.7**	0.28***	1.12***	4247.8**	0.32***	0.036***
	(1820.5)	(0.093)	(0.32)	(1759.9)	(0.094)	(0.013)
Wald tests (F)						
Female vs. Male	0.678	0.750	0.072	0.603	0.260	0.245
p-value	0.411	0.388	0.788	0.438	0.608	0.621
By Sex of Head						
Female Headed*Program Area	2414.2**	0.53***	0.52***	2477.3**	0.47***	0.100*
	(945.5)	(0.12)	(0.12)	(1016.7)	(0.13)	(0.057)
Male Headed*Program Area	3762.5***	0.30***	0.31***	3693.3***	0.34***	0.029***
	(1085.3)	(0.068)	(0.069)	(1050.8)	(0.070)	(0.0088)
Wald tests (F)						
Female Headed vs. Male Headed	0.967	2.600	0.158	0.736	0.750	1.565
p-value	0.326	0.107	0.691	0.391	0.387	0.212
By Respondent Education						
Over 7 years*Program Area	4242.4***	0.38***	1.18***	4191.5***	0.39***	0.056***
	(1245.8)	(0.073)	(0.23)	(1218.2)	(0.075)	(0.021)
Under 8 years*Program Area	1878.9*	0.26***	0.83***	1848.3*	0.29***	0.020**

	Sale of AF Products Ksh (2016)	Sale of AF over Ksh 1000 (2016)	Sale of AF Products (log) (2016)	Sale of AF Products Ksh (difference)	Sale of AF over Ksh 1000 (difference)	Sale of AF Products (log) (difference)
	(972.3)	(0.087)	(0.23)	(965.8)	(0.088)	(0.010)
Wald tests (F)						
Under 8s vs. Over 7s	2.277	1.180	1.323	2.227	0.970	2.348
p-value	0.132	0.278	0.251	0.136	0.325	0.126
By Baseline Wealth Status (Asset-based)						
Asset Poor*Program Area	2462.8***	0.43***	1.20***	2433.7***	0.45***	0.021***
	(770.1)	(0.089)	(0.24)	(763.5)	(0.089)	(0.0066)
Asset Rich*Program Area	4309.5***	0.26***	0.90***	4248.5***	0.28***	0.066**
	(1529.4)	(0.078)	(0.24)	(1489.1)	(0.077)	(0.028)
Wald tests (F)						
Asset Rich vs. Asset Poor	1.232	2.270	0.903	1.219	2.340	2.497
p-value	0.268	0.132	0.342	0.270	0.126	0.115
By Landholding Size						
< 2 acre*Program Area	2011.7***	0.37***	0.92***	1996.1***	0.38***	0.017***
	(748.3)	(0.087)	(0.21)	(749.6)	(0.089)	(0.0061)
≥ 2 acre*Program Area	4751.3***	0.31***	1.16***	4653.0***	0.33***	0.072**
	(1529.2)	(0.080)	(0.26)	(1473.9)	(0.079)	(0.029)
Wald tests (F)						
Under 2 vs. 2 or overs	2.907	0.270	0.582	2.876	0.160	3.476
p-value	0.089	0.603	0.446	0.091	0.687	0.063
By Dairy Producer						
Dairy*Program Area	3665.4**	0.18**	0.66**	3656.5**	0.21**	0.069**
	(1483.3)	(0.081)	(0.26)	(1474.5)	(0.081)	(0.031)
No Dairy*Program Area	3240.2***	0.50***	1.38***	3165.1***	0.51***	0.022***
	(877.4)	(0.080)	(0.21)	(860.8)	(0.082)	(0.0086)
Wald tests (F)						
Dairy vs. No dairy	0.071	8.940	5.744	0.091	8.420	2.106
p-value	0.790	0.003	0.017	0.763	0.004	0.147
By Official Position						
Official Position*Program Area	3163.4**	0.28***	0.93***	3097.4**	0.31***	0.042*
	(1349.0)	(0.075)	(0.23)	(1311.6)	(0.076)	(0.023)
No Official Position*Program Area	3651.4***	0.39***	1.15***	3629.3***	0.40***	0.044***
	(1020.5)	(0.085)	(0.22)	(1015.4)	(0.083)	(0.017)
Wald tests (F)						
Off. Pos. vs. No Off. Pos.	0.096	1.030	0.602	0.114	0.910	0.003
p-value	0.757	0.311	0.438	0.736	0.340	0.960

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with programme area (p=<0.1) used in all models; VSZ dummies used for fixed effects in all linear models

A5.6: Firewood Cash Value and Collection Time, Differential Effects

	Cash Value of Firewood from Farm				*Hours in Month Collecting Firewood*			
	2016	**2016 log**	**Dif.**	**Dif.(log)**	**2016**	**2016 log**	**Dif.**	**Dif.(log)**
Overall OLS ITT estimates	143.8***	0.54***	138.4***	0.39***	-1.47***	-0.12***	-0.87	-0.087**
	(40.3)	(0.12)	(39.7)	(0.13)	(0.42)	(0.029)	(0.60)	(0.034)
Observations	**2790**	**2790**	**2790**	**2790**	**2732**	**2732**	**2732**	**2732**
By Village Sampling Zone								
Sirisia/Malakisi*Program Area	189.3***	0.72***	178.3***	0.67***	-1.80**	-0.084*	-1.69*	-0.063
	(68.4)	(0.20)	(67.3)	(0.22)	(0.76)	(0.049)	(1.01)	(0.060)
Bumula*Program Area	300.3***	1.06***	293.5***	0.84***	-1.38	-0.12*	-0.12	-0.051
	(86.4)	(0.23)	(84.8)	(0.26)	(0.88)	(0.060)	(1.16)	(0.067)
Kimilili/Ndivisi*Program Area	9.71	-0.21	13.4	-0.16	0.75	0.0035	1.35	0.020
	(85.6)	(0.23)	(86.9)	(0.26)	(0.88)	(0.066)	(1.32)	(0.076)
Tongaren/Likuyani*Program Area	58.3	0.48*	51.7	0.14	-3.22***	-0.26***	-2.79**	-0.24***
	(80.7)	(0.26)	(78.8)	(0.26)	(0.82)	(0.057)	(1.30)	(0.067)
Wald tests (F)								
Sirisia/Malakisi vs. Bumula	1.000	1.206	1.113	0.236	0.131	0.183	1.049	0.016
p-value	0.318	0.273	0.292	0.627	0.717	0.669	0.306	0.900
Sirisia/Malakisi vs. Kimilili/Ndivisi	2.684	9.205	2.237	5.797	4.546	1.152	3.355	0.732
p-value	0.102	0.003	0.135	0.016	0.034	0.284	0.068	0.393
Sirisia/Malakisi vs. Tongaren/Likuyani	1.534	0.542	1.493	2.477	1.612	5.694	0.459	3.988
p-value	0.216	0.462	0.222	0.116	0.205	0.017	0.499	0.046
Bumula vs. Kimilili/Ndivisi	5.623	15.400	5.219	7.368	3.051	1.937	0.697	0.510
p-value	0.018	0.000	0.023	0.007	0.081	0.165	0.404	0.476
Bumula vs. Tongaren/Likuyani	4.068	2.853	4.242	3.743	2.297	3.076	2.342	4.050
p-value	0.044	0.092	0.040	0.054	0.130	0.080	0.127	0.045
Kimilili/Ndivisi vs. Tongaren/Likuyani	0.176	3.964	0.110	0.685	10.577	9.154	4.852	6.564
p-value	0.675	0.047	0.741	0.408	0.001	0.003	0.028	0.011
By Gender								
Female*Program Area	174.9***	0.67***	166.4***	0.43***	-1.19**	-0.075**	0.11	-0.020
	(47.5)	(0.16)	(46.7)	(0.16)	(0.57)	(0.038)	(0.78)	(0.044)
Male*Program Area	95.0	0.32**	94.4	0.32	-1.91***	-0.19***	-2.39***	-0.19***
	(66.5)	(0.16)	(67.3)	(0.20)	(0.60)	(0.043)	(0.82)	(0.050)
Wald tests (F)								
Female vs. Male	1.025	2.763	0.813	0.220	0.718	3.998	5.419	7.287
p-value	0.312	0.097	0.368	0.639	0.397	0.046	0.020	0.007
By Sex of Head								
Female Headed*Program Area	123.2*	0.50**	131.2*	0.42*	-1.71*	-0.11*	-0.32	-0.045
	(69.9)	(0.22)	(69.5)	(0.23)	(0.88)	(0.063)	(1.19)	(0.074)
Male Headed*Program Area	152.2***	0.55***	143.5***	0.39***	-1.39***	-0.12***	-1.04	-0.099***
	(45.2)	(0.13)	(44.9)	(0.15)	(0.47)	(0.032)	(0.67)	(0.036)
Wald tests (F)								
Female Headed vs. Male Headed	0.139	0.039	0.025	0.016	0.106	0.043	0.290	0.473
p-value	0.709	0.843	0.875	0.900	0.745	0.835	0.591	0.492
By Respondent Education								
Over 7 years*Program Area	140.8***	0.59***	142.0***	0.51***	-1.44***	-0.14***	-0.56	-0.10**
	(49.6)	(0.13)	(49.4)	(0.15)	(0.49)	(0.036)	(0.71)	(0.042)
Under 8 years*Program Area	151.3**	0.44**	133.5**	0.17	-1.49*	-0.083*	-1.43	-0.063
	(62.2)	(0.18)	(62.3)	(0.20)	(0.77)	(0.047)	(1.01)	(0.055)
Wald tests (F)								

	Cash Value of Firewood from Farm				*Hours in Month Collecting Firewood*			
	2016	**2016 log**	**Dif.**	**Dif.(log)**	**2016**	**2016 log**	**Dif.**	**Dif.(log)**
Under 8s vs. Over 7s	0.019	0.565	0.012	2.054	0.003	0.832	0.524	0.299
p-value	0.891	0.453	0.913	0.153	0.953	0.362	0.469	0.585
By Baseline Wealth Status (Asset-based)								
Asset Poor*Program Area	172.6***	0.68***	160.7***	0.54***	-1.63**	-0.11***	-0.74	-0.054
	(48.8)	(0.17)	(48.3)	(0.18)	(0.63)	(0.041)	(0.83)	(0.047)
Asset Rich*Program Area	107.2*	0.38***	109.8*	0.24	-1.28**	-0.12***	-1.01	-0.12***
	(58.7)	(0.14)	(58.4)	(0.16)	(0.54)	(0.038)	(0.74)	(0.044)
Wald tests (F)								
Asset Rich vs. Asset Poor	0.822	2.187	0.501	1.729	0.173	0.023	0.074	1.256
p-value	0.365	0.140	0.479	0.189	0.677	0.880	0.786	0.263
By Landholding Size								
< 2 acre*Program Area	162.4***	0.63***	144.2***	0.43**	-1.36**	-0.12***	-0.82	-0.087*
	(53.0)	(0.19)	(52.5)	(0.20)	(0.62)	(0.041)	(0.85)	(0.046)
≥ 2 acre*Program Area	111.9**	0.40***	122.4**	0.35**	-1.52***	-0.12***	-0.90	-0.086*
	(55.7)	(0.13)	(56.0)	(0.16)	(0.55)	(0.040)	(0.79)	(0.047)
Wald tests (F)								
Under 2 vs. 2 or overs	0.476	1.059	0.087	0.099	0.039	0.002	0.006	0.000
p-value	0.491	0.304	0.768	0.753	0.844	0.965	0.938	0.997
By Dairy Producer								
Dairy*Program Area	71.4	0.35**	67.9	0.19	-1.11**	-0.093**	-0.25	-0.042
	(57.5)	(0.15)	(56.9)	(0.16)	(0.55)	(0.038)	(0.69)	(0.042)
No Dairy*Program Area	201.4***	0.69***	194.0***	0.55***	-1.76***	-0.14***	-1.37*	-0.12**
	(51.1)	(0.16)	(51.4)	(0.18)	(0.55)	(0.039)	(0.82)	(0.048)
Wald tests (F)								
Dairy vs. No dairy	3.221	2.955	2.974	2.495	0.806	0.897	1.355	1.842
p-value	0.073	0.086	0.085	0.115	0.370	0.344	0.245	0.175
By Official Position								
Official Position*Program Area	94.4	0.47***	98.7	0.36**	-1.89***	-0.16***	-1.74**	-0.14***
	(60.1)	(0.15)	(60.1)	(0.18)	(0.55)	(0.040)	(0.84)	(0.048)
No Official Position*Program Area	185.3***	0.59***	171.1***	0.41**	-1.00*	-0.075*	0.010	-0.036
	(49.5)	(0.16)	(49.2)	(0.17)	(0.60)	(0.041)	(0.81)	(0.048)
Wald tests (F)								
Off. Pos. vs. No Off. Pos.	1.498	0.376	0.931	0.049	1.276	2.299	2.356	2.182
p-value	0.222	0.540	0.335	0.824	0.259	0.130	0.126	0.140

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with programme area (p=<0.1) used in all models; VSZ dummies used for fixed effects in all linear models

A5.7: HH Consumption Expenditure & Asset Indices Weighted by these Data, Differential Effects

	2016 HH consumption expenditure (CE)/capita				***Asset indices weighted by 2016 CE data***			
	16 CE w/ covariates only	16 CE log w covariates only	16 CE w/ CE 07 asset index	16 CE log w/ CE 07 asset index (log)	16 CE asset weighted w/ CE 07 asset weights	16 CE asset weighted w/ CE 07 asset weights (log)	Differenced 16 CE asset weighted	Differenced 16 CE asset weighted (log)
Overall OLS ITT estimates	0.079	0.032	0.067	0.032	0.10*	0.017	0.10*	0.017
	(0.12)	(0.024)	(0.100)	(0.020)	(0.060)	(0.011)	(0.062)	(0.011)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
By Village Sampling Zone								
Sirisia/Malakisi*Program Area	0.20	0.076**	0.30*	0.092***	0.011	0.0092	-0.0011	0.012
	(0.18)	(0.037)	(0.16)	(0.031)	(0.12)	(0.024)	(0.12)	(0.024)
Bumula*Program Area	0.17	0.040	0.16	0.045	0.023	-0.0044	0.023	-0.0035
	(0.25)	(0.048)	(0.22)	(0.043)	(0.098)	(0.021)	(0.099)	(0.021)
Kimilili/Ndivisi*Program Area	-0.34	-0.066	-0.35*	-0.063	0.079	0.012	0.080	0.012
	(0.23)	(0.050)	(0.20)	(0.045)	(0.13)	(0.021)	(0.13)	(0.021)
Tongaren/Likuyani*Program Area	0.24	0.067	0.11	0.045	0.29**	0.052**	0.31**	0.048**
	(0.31)	(0.057)	(0.22)	(0.043)	(0.12)	(0.021)	(0.13)	(0.021)
Wald tests (F)								
Sirisia/Malakisi vs. Bumula	0.012	0.343	0.260	0.792	0.005	0.186	0.025	0.251
p-value	0.914	0.558	0.610	0.374	0.941	0.666	0.876	0.616
Sirisia/Malakisi vs. Kimilili/Ndivisi	3.323	5.203	6.436	8.161	0.146	0.006	0.201	0.000
p-value	0.069	0.023	0.012	0.004	0.703	0.939	0.654	0.999
Sirisia/Mal.i vs. Tongaren/Lik.	0.099	0.016	0.523	0.806	2.612	1.800	2.955	1.283
p-value	0.753	0.901	0.470	0.370	0.107	0.180	0.086	0.258
Bumula vs. Kimilili/Ndivisi	1.718	2.334	2.888	2.988	0.117	0.294	0.116	0.288
p-value	0.191	0.127	0.090	0.085	0.732	0.588	0.733	0.592
Bumula vs. Tongaren/Likuyani	0.022	0.139	0.031	0.000	2.832	3.728	2.978	3.130
p-value	0.881	0.710	0.860	0.994	0.093	0.054	0.085	0.078
Kimilili/Ndv. vs. Tongaren/Lik.	1.115	2.947	2.162	2.815	1.319	1.767	1.443	1.431
p-value	0.292	0.087	0.142	0.094	0.251	0.184	0.230	0.232
By Gender – OLS								
Female*Program Area	0.13	0.046	0.11	0.045*	0.11	0.026*	0.11	0.025*
	(0.15)	(0.030)	(0.12)	(0.026)	(0.069)	(0.013)	(0.071)	(0.013)
Male*Program Area	0.0023	0.0093	0.0070	0.012	0.094	0.0043	0.094	0.0049
	(0.20)	(0.036)	(0.17)	(0.032)	(0.097)	(0.016)	(0.098)	(0.017)
Wald tests (F)								
Female vs. Male	0.255	0.662	0.204	0.649	0.009	1.163	0.016	1.028
p-value	0.614	0.416	0.652	0.421	0.923	0.281	0.899	0.311
By Gender – Robust Regression								
Female*Program Area	0.18**	0.048*	0.15*	0.044**	0.13**	0.031**	0.13**	0.030**
	(0.091)	(0.025)	(0.082)	(0.022)	(0.054)	(0.012)	(0.054)	(0.012)
Male*Program Area	0.077	0.011	0.095	0.013	-0.048	0.00070	-0.047	-0.00056
	(0.11)	(0.031)	(0.10)	(0.027)	(0.068)	(0.015)	(0.068)	(0.016)
Wald tests (F)								
Female vs. Male	0.535	0.919	0.208	0.837	4.391	2.344	4.397	2.305
p-value	0.465	0.338	0.648	0.360	0.036	0.126	0.036	0.129
By Gender – Quantile Regression								
Female*Program Area	0.18	0.043	0.21*	0.049*	0.17**	0.054***	0.18***	0.042***
	(0.12)	(0.029)	(0.11)	(0.027)	(0.072)	(0.015)	(0.068)	(0.015)
Male*Program Area	0.035	0.0057	0.067	0.036	-0.046	-0.012	-0.050	-0.021
	(0.15)	(0.036)	(0.14)	(0.033)	(0.090)	(0.019)	(0.085)	(0.019)

	2016 HH consumption expenditure (CE)/capita				***Asset indices weighted by 2016 CE data***			
	16 CE w/ covariates only	16 CE log w covariates only	16 CE w/ CE 07 asset index	16 CE log w/ CE 07 asset index (log)	16 CE asset weighted w/ CE 07 asset weights	16 CE asset weighted w/ CE 07 asset weights (log)	Differenced 16 CE asset weighted	Differenced 16 CE asset weighted (log)
Wald tests (F)								
Female vs. Male	0.616	0.660	0.612	0.105	3.556	7.742	4.462	6.786
p-value	0.432	0.417	0.434	0.746	0.059	0.005	0.035	0.009
By Sex of Head								
Female Headed*Program Area	0.23	0.052	0.11	0.052	0.036	0.016	0.051	0.014
	(0.23)	(0.047)	(0.20)	(0.047)	(0.12)	(0.022)	(0.12)	(0.022)
Male Headed*Program Area	0.022	0.023	0.052	0.023	0.12*	0.018	0.12*	0.019
	(0.13)	(0.025)	(0.11)	(0.025)	(0.065)	(0.012)	(0.066)	(0.012)
Wald tests (F)								
Female Headed vs. Male Headed	0.793	0.334	0.074	0.037	0.419	0.005	0.254	0.052
p-value	0.374	0.564	0.786	0.848	0.518	0.943	0.614	0.819
By Respondent Education								
Under 8 years*Program Area	0.014	0.020	0.0043	0.020	0.15*	0.023*	0.16*	0.023*
	(0.15)	(0.027)	(0.12)	(0.023)	(0.079)	(0.013)	(0.080)	(0.014)
Over 7 years*Program Area	0.15	0.042	0.17	0.051*	-0.016	0.0027	-0.019	0.0045
	(0.16)	(0.036)	(0.13)	(0.031)	(0.080)	(0.016)	(0.081)	(0.016)
Wald tests (F)								
Under 8s vs. Over 7s	0.494	0.303	1.006	0.844	2.470	0.973	2.537	0.783
p-value	0.483	0.582	0.316	0.359	0.117	0.324	0.112	0.377
By Baseline Wealth Status (Asset-based)								
Asset Poor*Program Area	0.13	0.033	0.16	0.038	0.065	0.015	0.061	0.016
	(0.12)	(0.028)	(0.11)	(0.026)	(0.070)	(0.016)	(0.071)	(0.016)
Asset Rich*Program Area	-0.045	0.015	-0.037	0.023	0.13	0.019	0.13	0.020
	(0.19)	(0.030)	(0.15)	(0.026)	(0.093)	(0.014)	(0.095)	(0.014)
Wald tests (F)								
Asset Rich vs. Asset Poor	0.711	0.223	1.245	0.198	0.385	0.039	0.407	0.052
p-value	0.400	0.637	0.265	0.656	0.535	0.844	0.524	0.819
By Landholding Size								
< 2 acre*Program Area	0.25*	0.059*	0.11	0.034	0.13*	0.023	0.14*	0.018
	(0.15)	(0.030)	(0.13)	(0.027)	(0.073)	(0.015)	(0.074)	(0.015)
≥ 2 acre*Program Area	-0.14	-0.0051	0.012	0.027	0.070	0.010	0.051	0.016
	(0.18)	(0.032)	(0.14)	(0.026)	(0.091)	(0.014)	(0.091)	(0.015)
Wald tests (F)								
Under 2 vs. 2 or overs	3.197	2.599	0.282	0.049	0.274	0.455	0.739	0.013
p-value	0.074	0.108	0.596	0.825	0.601	0.501	0.391	0.910
By Dairy Producer								
Dairy*Program Area	-0.17	-0.012	-0.11	-0.0034	0.044	0.010	0.037	0.012
	(0.19)	(0.033)	(0.15)	(0.028)	(0.080)	(0.014)	(0.082)	(0.014)
No Dairy*Program Area	0.28*	0.066**	0.20*	0.060**	0.15*	0.023	0.15*	0.022
	(0.15)	(0.030)	(0.12)	(0.026)	(0.079)	(0.015)	(0.080)	(0.015)
Wald tests (F)								
Dairy vs. No dairy	4.136	3.410	2.951	3.245	0.927	0.474	1.220	0.292
p-value	0.043	0.065	0.087	0.072	0.336	0.491	0.270	0.589
By Dairy Producer – Robust Regression								
Dairy*Program Area	-0.0023	-0.011	-0.035	-0.0099	0.065	0.015	0.065	0.015
	(0.11)	(0.029)	(0.096)	(0.025)	(0.063)	(0.014)	(0.063)	(0.015)
No Dairy*Program Area	0.23**	0.064**	0.24***	0.065***	0.066	0.022*	0.066	0.020
	(0.094)	(0.026)	(0.086)	(0.022)	(0.056)	(0.013)	(0.056)	(0.013)

	2016 HH consumption expenditure (CE)/capita				***Asset indices weighted by 2016 CE data***			
	16 CE w/ covariates only	16 CE log w covariates only	16 CE w/ CE 07 asset index	16 CE log w/ CE 07 asset index (log)	16 CE asset weighted w/ CE 07 asset weights	16 CE asset weighted w/ CE 07 asset weights (log)	Differenced 16 CE asset weighted	Differenced 16 CE asset weighted (log)
Wald tests (F)								
Dairy vs. No dairy	2.605	3.797	4.693	5.035	0.000	0.146	0.000	0.059
p-value	0.107	0.051	0.030	0.025	0.990	0.703	0.994	0.807
By Dairy Producer – Quantile Regression								
Dairy*Program Area	-0.10	-0.028	-0.063	0.00044	-0.00083	0.0075	0.024	0.014
	(0.14)	(0.034)	(0.13)	(0.032)	(0.083)	(0.018)	(0.077)	(0.019)
No Dairy*Program Area	0.24**	0.061**	0.31***	0.077***	0.13*	0.027*	0.13*	0.033**
	(0.12)	(0.030)	(0.12)	(0.028)	(0.074)	(0.016)	(0.068)	(0.017)
Wald tests (F)								
Dairy vs. No dairy	3.591	3.869	4.427	3.353	1.421	0.706	1.021	0.592
p-value	0.058	0.049	0.035	0.067	0.233	0.401	0.312	0.442
By Official Position								
Official Position*Program Area	0.13	0.047	0.17	0.053**	0.18**	0.023	0.18**	0.024
	(0.18)	(0.032)	(0.15)	(0.026)	(0.088)	(0.015)	(0.090)	(0.015)
No Official Position*Program Area	-0.0093	0.0097	-0.049	0.0076	0.012	0.0095	0.017	0.0091
	(0.14)	(0.029)	(0.12)	(0.026)	(0.070)	(0.013)	(0.071)	(0.013)
Wald tests (F)								
Off. Pos. vs. No Off. Pos.	0.469	0.949	1.473	1.952	2.796	0.477	2.402	0.605
p-value	0.494	0.331	0.226	0.163	0.095	0.490	0.122	0.437

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with programme area (p=<0.1) used in all models; VSZ dummies used for fixed effects in all linear models

A5.8: Household Asset Wealth and Accumulation, Differential Effects

	Principal Component Asset Indices				***Raw Asset Positive Gains***				
	2016 Overall	2016 by VSZ	Dif. Overall	Dif. by VSZ	Overall	House Char.	Home Durables	Productive Assets	Livestock
Overall OLS ITT estimates	0.050*	0.040	0.069***	0.048**	0.59***	0.16**	0.17*	0.14**	0.12*
	(0.026)	(0.028)	(0.025)	(0.023)	(0.21)	(0.071)	(0.10)	(0.070)	(0.068)
Observations	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
By Village Sampling Zone									
Sirisia/Malakisi*Program Area	0.025	0.013	0.035	0.031	0.40	0.032	0.15	0.076	0.15
	(0.053)	(0.056)	(0.046)	(0.049)	(0.44)	(0.12)	(0.20)	(0.14)	(0.15)
Bumula*Program Area	0.037	0.031	0.051	0.057	0.80*	0.098	0.27	0.20	0.23*
	(0.056)	(0.060)	(0.051)	(0.049)	(0.43)	(0.15)	(0.22)	(0.13)	(0.12)
Kimilili/Ndivisi*Program Area	-0.013	-0.020	0.020	0.020	0.0017	0.26*	-0.21	0.0071	-0.042
	(0.052)	(0.056)	(0.049)	(0.048)	(0.42)	(0.14)	(0.21)	(0.15)	(0.14)
Tongaren/Likuyani*Program Area	0.14***	0.13**	0.16***	0.078**	1.08***	0.26	0.44**	0.27*	0.12
	(0.053)	(0.051)	(0.055)	(0.038)	(0.42)	(0.16)	(0.19)	(0.14)	(0.13)
Wald tests (F)									
Sirisia/Malakisi vs. Bumula	0.028	0.050	0.057	0.144	0.433	0.126	0.173	0.441	0.196
p-value	0.867	0.823	0.811	0.704	0.511	0.723	0.677	0.507	0.658
Sirisia/Malakisi vs. Kimilili/Ndivisi	0.251	0.168	0.047	0.026	0.436	1.506	1.573	0.110	0.820
p-value	0.617	0.682	0.828	0.873	0.509	0.220	0.211	0.740	0.366
Sirisia/Malakisi vs. Tongaren/Likuyani	2.631	2.381	3.200	0.580	1.288	1.373	1.167	0.931	0.020
p-value	0.106	0.124	0.074	0.447	0.257	0.242	0.281	0.335	0.887
Bumula vs. Kimilili/Ndivisi	0.427	0.380	0.192	0.291	1.737	0.619	2.641	0.932	2.100
p-value	0.514	0.538	0.662	0.590	0.188	0.432	0.105	0.335	0.148
Bumula vs. Tongaren/Likuyani	1.922	1.537	2.191	0.112	0.220	0.579	0.366	0.117	0.424
p-value	0.166	0.216	0.140	0.738	0.639	0.447	0.546	0.733	0.515
Kim./Ndv. vs. Tong./Likuyani	4.252	3.615	3.509	0.876	3.202	0.001	5.181	1.557	0.727
p-value	0.040	0.058	0.062	0.350	0.074	0.973	0.023	0.213	0.394
By Gender									
Female*Program Area	0.100***	0.091***	0.098***	0.072***	0.95***	0.23***	0.26**	0.26***	0.20***
	(0.029)	(0.031)	(0.028)	(0.026)	(0.24)	(0.084)	(0.12)	(0.077)	(0.073)
Male*Program Area	-0.027	-0.039	0.022	0.010	0.025	0.046	0.038	-0.042	-0.014
	(0.045)	(0.047)	(0.043)	(0.041)	(0.36)	(0.13)	(0.16)	(0.12)	(0.12)
Wald tests (F)									
Female vs. Male	6.305	5.841	2.475	1.734	5.077	1.548	1.247	5.364	2.702
p-value	0.012	0.016	0.116	0.189	0.025	0.214	0.265	0.021	0.101
By Sex of Head									
Female Headed*Program Area	0.094**	0.079	0.072*	0.063	0.78**	0.094	0.36**	0.30***	0.033
	(0.047)	(0.049)	(0.042)	(0.039)	(0.36)	(0.14)	(0.17)	(0.12)	(0.11)
Male Headed*Program Area	0.038	0.030	0.067**	0.043	0.54**	0.17**	0.12	0.10	0.15*
	(0.030)	(0.031)	(0.029)	(0.027)	(0.24)	(0.084)	(0.11)	(0.080)	(0.080)
Wald tests (F)									
Female Headed vs. Male Headed	1.067	0.799	0.010	0.200	0.347	0.252	1.532	2.273	0.691
p-value	0.302	0.372	0.920	0.655	0.556	0.616	0.216	0.132	0.406
By Respondent Education									
Over 7 years*Program Area	0.044	0.036	0.079**	0.060*	0.68**	0.18*	0.14	0.17*	0.044
	(0.035)	(0.036)	(0.034)	(0.031)	(0.27)	(0.099)	(0.12)	(0.091)	(0.035)
Under 8 years*Program Area	0.054	0.039	0.042	0.019	0.36	0.11	0.20	0.074	0.054
	(0.033)	(0.035)	(0.030)	(0.029)	(0.28)	(0.091)	(0.15)	(0.089)	(0.033)
Wald tests (F)									
Under 8s vs. Over 7s	0.041	0.004	0.742	1.003	0.725	0.266	0.112	0.636	3.205

	Principal Component Asset Indices				***Raw Asset Positive Gains***				
	2016 Overall	2016 by VSZ	Dif. Overall	Dif. by VSZ	Overall	House Char.	Home Durables	Productive Assets	Livestock
p-value	0.840	0.950	0.389	0.317	0.395	0.607	0.739	0.426	0.074
By Baseline Wealth Status (Asset-based)									
Asset Poor*Program Area	0.053*	0.043	0.053**	0.043	0.60**	0.12	0.24*	0.10	0.10
	(0.030)	(0.032)	(0.026)	(0.027)	(0.26)	(0.075)	(0.14)	(0.091)	(0.091)
Asset Rich*Program Area	0.048	0.037	0.070*	0.043	0.53*	0.17	0.088	0.19**	0.19**
	(0.039)	(0.041)	(0.037)	(0.033)	(0.28)	(0.11)	(0.14)	(0.093)	(0.093)
Wald tests (F)									
Asset Rich vs. Asset Poor	0.011	0.014	0.161	0.000	0.043	0.183	0.591	0.506	0.277
p-value	0.917	0.907	0.688	0.990	0.835	0.669	0.442	0.477	0.599
By Landholding Size									
< 2 acre*Program Area	0.056*	0.045	0.069**	0.044	0.54**	0.17**	0.17	0.15	0.056
	(0.031)	(0.032)	(0.030)	(0.028)	(0.27)	(0.085)	(0.14)	(0.094)	(0.083)
≥ 2 acre*Program Area	0.044	0.035	0.062*	0.049	0.62**	0.13	0.19	0.14	0.17*
	(0.038)	(0.040)	(0.036)	(0.034)	(0.29)	(0.11)	(0.14)	(0.093)	(0.10)
Wald tests (F)									
Under 2 vs. 2 or overs	0.063	0.047	0.024	0.012	0.040	0.107	0.010	0.013	0.798
p-value	0.802	0.828	0.878	0.913	0.842	0.744	0.921	0.908	0.372
By Dairy Producer									
Dairy*Program Area	0.039	0.025	0.073**	0.051	0.56*	0.14	0.090	0.21**	0.13
	(0.036)	(0.038)	(0.035)	(0.034)	(0.29)	(0.10)	(0.14)	(0.092)	(0.100)
No Dairy*Program Area	0.060*	0.053	0.065**	0.045	0.62**	0.18**	0.24*	0.092	0.11
	(0.033)	(0.035)	(0.031)	(0.029)	(0.27)	(0.089)	(0.13)	(0.087)	(0.088)
Wald tests (F)									
Dairy vs. No dairy	0.201	0.725	0.035	0.020	0.030	0.112	0.682	1.081	0.022
p-value	0.654	0.395	0.852	0.887	0.864	0.739	0.409	0.299	0.883
By Official Position									
Official Position*Program Area	0.078**	0.066	0.094**	0.066*	0.78**	0.22**	0.15	0.22**	0.20**
	(0.039)	(0.041)	(0.038)	(0.034)	(0.30)	(0.11)	(0.15)	(0.090)	(0.092)
No Official Position*Program Area	0.019	0.0097	0.037	0.025	0.37	0.088	0.16	0.071	0.044
	(0.030)	(0.032)	(0.027)	(0.026)	(0.24)	(0.082)	(0.12)	(0.091)	(0.089)
Wald tests (F)									
Off. Pos. vs. No Off. Pos.	1.738	1.396	1.791	1.121	1.380	1.043	0.011	1.620	1.646
p-value	0.188	0.238	0.182	0.290	0.241	0.308	0.916	0.204	0.200
By Respondent Parcel Ownership									
Parcel Own*Program Area	-0.0036	-0.015	0.031	0.023	0.20	0.094	0.069	0.020	0.020
	(0.038)	(0.040)	(0.036)	(0.035)	(0.30)	(0.11)	(0.14)	(0.10)	(0.10)
No Parcel Own *Program Area	0.10***	0.096***	0.11***	0.073**	0.99***	0.23**	0.28**	0.27***	0.22***
	(0.032)	(0.034)	(0.032)	(0.030)	(0.27)	(0.094)	(0.14)	(0.084)	(0.080)
Wald tests (F)									
Par.Own vs. No Par. Own	5.145	4.821	2.647	1.241	4.071	0.872	4.170	2.696	5.145
p-value	0.024	0.029	0.104	0.266	0.044	0.351	0.042	0.101	0.024

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with programme area (p=<0.1) used in all models; VSZ dummies used for fixed effects in all linear models

A5.9: Coping Strategy Index & Food Security Measures, Differential Effects

	Coping Strategies			***MDD-W***			***Food Shortage Months***		
	Full Score	> 2 points	Asset Sale	Score/10	>4 points	Meat	Last 12 months	Dif. from baseline	Reduction from baseline
Overall OLS ITT estimates	-0.090	-0.073	-0.25**	0.15	0.11*	0.060	-0.088	-0.055	-0.014
	(0.16)	(0.068)	(0.12)	(0.096)	(0.063)	(0.059)	(0.089)	(0.074)	(0.055)
Observations	**2018**	**2018**	**2018**	**2790**	**2790**	**2790**	**2790**	**2790**	**2790**
By Village Sampling Zone									
Sirisia/Malakisi*Program Area	-0.25	-0.21*	-0.084	0.053	0.18*	-0.028	-0.16	-0.14	0.041
	(0.23)	(0.12)	(0.21)	(0.14)	(0.10)	(0.099)	(0.17)	(0.16)	(0.038)
Bumula*Program Area	-0.029	0.033	-0.47**	0.028	0.10	0.082	-0.19	-0.24	-0.0048
	(0.27)	(0.13)	(0.24)	(0.16)	(0.12)	(0.10)	(0.18)	(0.15)	(0.039)
Kimilili/Ndivisi*Program Area	0.43	0.064	-0.027	0.30	0.042	0.20	-0.12	0.11	-0.041
	(0.37)	(0.15)	(0.26)	(0.24)	(0.13)	(0.15)	(0.18)	(0.15)	(0.041)
Tongaren/Likuyani*Program Area	-0.43	-0.17	-0.35	0.23	0.11	0.0071	0.12	0.068	-0.018
	(0.42)	(0.16)	(0.22)	(0.23)	(0.14)	(0.12)	(0.17)	(0.13)	(0.040)
Wald tests (F)									
Sirisia/Malakisi vs. Bumula	0.395	1.94	1.53	0.013	0.240	0.590	0.014	0.185	0.680
p-value	0.530	0.163	0.217	0.908	0.626	0.441	0.905	0.667	0.408
Sir. /Malakisi vs. Kim./Ndivisi	2.477	1.97	0.03	0.770	0.660	1.520	0.027	1.398	2.200
p-value	0.112	0.161	0.86	0.381	0.418	0.217	0.870	0.238	0.138
Sir./Malakisi vs. Ton./Likuyani	0.137	0.04	0.76	0.466	0.180	0.050	1.424	1.021	1.170
p-value	0.712	0.837	0.384	0.495	0.675	0.822	0.233	0.313	0.280
Bumula vs. Kimilili/Ndivisi	1.000	0.02	1.64	0.850	0.110	0.380	0.075	2.644	0.410
p-value	0.318	0.876	0.200	0.357	0.740	0.540	0.784	0.105	0.522
Bumula vs. Tongaren/Likuyani	0.650	0.92	0.14	0.547	0.000	0.230	1.707	2.293	0.060
p-value	0.421	0.338	0.703	0.460	0.985	0.632	0.192	0.131	0.804
Kim./Ndivisi vs. Ton./Likuyani	2.353	1.09	0.90	0.036	0.100	0.910	1.015	0.046	0.140
p-value	0.126	0.297	0.344	0.849	0.750	0.341	0.314	0.830	0.707
By Gender									
Female*Program Area	-0.25	-0.13	-0.34**	0.22*	0.099	0.11	-0.083	0.016	-0.034
	(0.23)	(0.089)	(0.14)	(0.12)	(0.078)	(0.078)	(0.11)	(0.094)	(0.072)
Male*Program Area	0.14	0.016	0.0017	0.033	0.12	-0.011	-0.097	-0.17	0.015
	(0.20)	(0.11)	(0.20)	(0.14)	(0.10)	(0.092)	(0.13)	(0.12)	(0.087)
Wald tests (F)									
Female vs. Male	1.71	1.08	2.22	1.061	0.030	0.920	0.008	1.517	0.010
p-value	0.1911	0.299	0.136	0.303	0.856	0.336	0.928	0.219	0.663
By Sex of Head									
Female Headed*Program Area	-0.33	-0.22	-0.39**	-0.039	-0.093	0.026	-0.077	0.11	-0.071
	(0.33)	(0.14)	(0.19)	(0.18)	(0.12)	(0.12)	(0.15)	(0.15)	(0.11)
Male Headed*Program Area	-0.032	-0.035	-0.20	0.21**	0.17**	0.071	-0.088	-0.10	0.0028
	(0.17)	(0.073)	(0.12)	(0.10)	(0.068)	(0.064)	(0.10)	(0.081)	(0.060)
Wald tests (F)									
F. Headed vs. Male Headed	0.80	1.59	1.02	1.647	4.150	0.120	0.004	1.738	0.400
p-value	0.371	0.207	0.312	0.200	0.042	0.730	0.948	0.188	0.527
By Respondent Education									
Over 7 years*Program Area	-0.053	-0.077	-0.16	0.16	0.16**	0.027	-0.18*	-0.13	-0.0026
	(0.18)	(0.080)	(0.13)	(0.11)	(0.074)	(0.067)	(0.10)	(0.089)	(0.024)

	Coping Strategies			***MDD-W***			***Food Shortage Months***		
	Full Score	> 2 points	Asset Sale	Score/10	>4 points	Meat	Last 12 months	Dif. from baseline	Reduction from baseline
Under 8 years*Program Area	-0.16	-0.066	-0.48***	0.13	0.011	0.12	0.098	0.094	-0.010
	(0.25)	(0.11)	(0.18)	(0.13)	(0.090)	(0.088)	(0.14)	(0.12)	(0.030)
Wald tests (F)									
Under 8s vs. Over 7s	0.15	0.01	3.10	0.073	2.160	0.850	3.057	2.388	0.040
p-value	0.698	0.923	0.079	0.787	0.142	0.357	0.081	0.123	0.850
By Baseline Wealth Status (Asset-based)									
Asset Poor*Program Area	-0.016	-0.068	-0.23	0.13	0.11	0.012	-0.026	-0.027	0.011
	(0.21)	(0.091)	(0.15)	(0.10)	(0.078)	(0.075)	(0.12)	(0.11)	(0.077)
Asset Rich*Program Area	-0.14	-0.065	-0.28*	0.14	0.081	0.091	-0.099	-0.077	-0.037
	(0.21)	(0.099)	(0.15)	(0.13)	(0.086)	(0.082)	(0.10)	(0.092)	(0.070)
Wald tests (F)									
Asset Rich vs. Asset Poor	0.19	0.00	0.08	0.002	0.100	0.560	0.263	0.135	0.240
p-value	0.660	0.986	0.773	0.961	0.757	0.453	0.609	0.714	0.627
By Landholding Size									
< 2 acre*Program Area	-0.11	-0.078	-0.25*	0.26**	0.13*	0.080	0.0040	0.027	-0.0086
	(0.22)	(0.096)	(0.14)	(0.12)	(0.077)	(0.076)	(0.11)	(0.10)	(0.072)
≥ 2 acre*Program Area	-0.059	-0.065	-0.26*	0.023	0.070	0.039	-0.17	-0.15	-0.019
	(0.20)	(0.090)	(0.15)	(0.12)	(0.082)	(0.075)	(0.11)	(0.095)	(0.074)
Wald tests (F)									
Under 2 vs. 2 or overs	0.03	0.01	0.000	2.918	0.440	0.190	1.632	1.822	0.010
p-value	0.857	0.921	0.985	0.088	0.506	0.665	0.202	0.178	0.916
By Dairy Producer									
Dairy*Program Area	0.0014	-0.016	-0.29**	0.090	0.14	0.039	-0.22*	-0.073	-0.077
	(0.19)	(0.093)	(0.15)	(0.13)	(0.088)	(0.082)	(0.12)	(0.10)	(0.083)
No Dairy*Program Area	-0.17	-0.12	-0.23	0.19*	0.083	0.077	0.017	-0.040	0.035
	(0.22)	(0.091)	(0.14)	(0.11)	(0.072)	(0.071)	(0.11)	(0.099)	(0.071)
Wald tests (F)									
Dairy vs. No dairy	0.43	0.70	0.12	0.542	0.320	0.150	2.416	0.054	1.120
p-value	0.511	0.403	0.732	0.462	0.570	0.696	0.121	0.817	0.289
By Official Position									
Official Position*Program Area	-0.12	-0.086	-0.24*	0.038	0.037	-0.020	-0.082	-0.040	-0.096
	(0.18)	(0.093)	(0.15)	(0.13)	(0.085)	(0.079)	(0.12)	(0.10)	(0.076)
No Official Position*Program Area	-0.065	-0.061	-0.26*	0.23**	0.16**	0.12*	-0.083	-0.062	0.059
	(0.23)	(0.092)	(0.14)	(0.11)	(0.077)	(0.073)	(0.11)	(0.094)	(0.072)
Wald tests (F)									
Off. Pos. vs. No Off. Pos.	0.04	0.04	0.01	1.812	1.420	2.210	0.000	0.028	0.000
p-value	0.850	0.836	0.920	0.179	0.234	0.137	0.995	0.866	0.115
By Parcel Ownership									
Parcel Own*Program Area	0.078	-0.049	-0.036	-0.039	0.059	-0.049	-0.070	-0.17*	0.0043
	(0.19)	(0.092)	(0.16)	(0.13)	(0.091)	(0.080)	(0.12)	(0.095)	(0.075)
No Parcel Own *Program Area	-0.29	-0.098	-0.41***	0.34***	0.15*	0.17**	-0.11	0.062	-0.033
	(0.25)	(0.097)	(0.15)	(0.13)	(0.085)	(0.086)	(0.12)	(0.11)	(0.080)
Wald tests (F)									
Par.Own vs. No Par. Own	1.499	0.140	3.490	4.716	0.600	3.570	0.052	2.674	0.120
p-value	0.222	0.712	0.062	0.030	0.438	0.059	0.820	0.103	0.725

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with programme area (p=<0.1) used in all models; VSZ dummies used for fixed effects in all linear models

A5.10: Education Progression & Spending & Economic Ladder, Differential Effects

	Secondary Progression for 14-17			***Education Expenditure per >5 Child***			***Economic Ladder***	
	% of children gone to sec.	Any child gone to sec.	All children gone to sec.	Ksh spent in previous year	Log of Ksh spent in previous year	Expenditure about median	Overall change	At least 1 step up
Overall OLS ITT estimates	-0.032	-0.16**	-0.070	-650.0	0.063	-0.040	0.039	0.027
	(0.022)	(0.077)	(0.070)	(1349.6)	(0.11)	(0.059)	(0.044)	(0.057)
Observations	**1693**	**1693**	**1693**	**2526**	**2526**	**2526**	**2790**	**2790**
By Village Sampling Zone								
Sirisia/Malakisi*Program Area	-0.038	-0.17	-0.10	887.4	0.20	-0.13	0.016	0.068
	(0.045)	(0.16)	(0.14)	(1874.0)	(0.22)	(0.11)	(0.091)	(0.11)
Bumula*Program Area	-0.13***	-0.52***	-0.37***	-2584.3	-0.17	-0.13	0.18*	0.15
	(0.037)	(0.13)	(0.13)	(3096.2)	(0.19)	(0.11)	(0.096)	(0.11)
Kimilili/Ndivisi*Program Area	0.068	0.083	0.30*	-5323.4*	-0.027	-0.098	-0.11	-0.19
	(0.048)	(0.16)	(0.16)	(2959.2)	(0.26)	(0.14)	(0.084)	(0.12)
Tongaren/Likuyani*Program Area	-0.011	-0.016	-0.041	3947.7	0.25	0.20	0.049	0.053
	(0.042)	(0.15)	(0.14)	(2612.9)	(0.19)	(0.12)	(0.072)	(0.11)
Wald tests (F)								
Sirisia/Malakisi vs. Bumula	2.383	2.820	1.890	0.930	1.663	0.000	1.591	0.280
p-value	0.123	0.093	0.169	0.336	0.198	0.996	0.208	0.599
Sirisia/Malakisi vs. Kimilili/Ndivisi	2.627	1.250	3.610	3.129	0.482	0.040	1.145	2.580
p-value	0.106	0.263	0.057	0.078	0.488	0.846	0.285	0.108
Sirisia/Mal.i vs. Tongaren/Lik.	0.194	0.480	0.090	0.916	0.027	4.180	0.082	0.010
p-value	0.660	0.486	0.760	0.339	0.871	0.041	0.775	0.925
Bumula vs. Kimilili/Ndivisi	10.083	8.010	10.080	0.404	0.191	0.040	5.494	4.150
p-value	0.002	0.005	0.002	0.525	0.662	0.850	0.020	0.042
Bumula vs. Tongaren/Likuyani	4.412	6.110	2.860	2.578	2.484	3.930	1.231	0.350
p-value	0.036	0.013	0.091	0.109	0.116	0.047	0.268	0.552
Kimilili/Ndv. vs. Tongaren/Lik.	1.497	0.190	2.580	5.606	0.723	2.640	2.113	2.070
p-value	0.222	0.660	0.108	0.018	0.396	0.104	0.147	0.151
By Gender – OLS								
Female*Program Area	-0.048*	-0.18**	-0.14	-893.1	0.0021	-0.042	0.11*	0.12*
	(0.026)	(0.092)	(0.084)	(1607.9)	(0.14)	(0.029)	(0.057)	(0.071)
Male*Program Area	-0.00100	-0.12	0.058	-247.7	0.16	0.029	-0.074	-0.13
	(0.036)	(0.13)	(0.12)	(2050.5)	(0.14)	(0.034)	(0.064)	(0.089)
Wald tests (F)								
Female vs. Male	1.172	0.140	2.090	0.071	0.689	2.670	5.042	5.070
p-value	0.280	0.703	0.149	0.790	0.407	0.102	0.025	0.024
By Sex of Head								
Female Headed*Program Area	0.0096	0.021	0.049	-1082.2	0.27	-0.042	0.12	0.14
	(0.046)	(0.16)	(0.15)	(2285.2)	(0.26)	(0.12)	(0.089)	(0.11)
Male Headed*Program Area	-0.043*	-0.22**	-0.10	14.9	-0.00014	-0.0100	0.020	-0.0052
	(0.024)	(0.087)	(0.080)	(1558.7)	(0.11)	(0.065)	(0.047)	(0.063)
Wald tests (F)								
Female Headed vs. Male Headed	1.022	1.650	0.770	0.240	0.934	0.080	1.073	1.610
p-value	0.313	0.200	0.380	0.624	0.334	0.773	0.301	0.205
By Respondent Education								
Over 7 years*Program Area	-0.042*	-0.18*	-0.13	-1413.1	0.038	-0.063	0.057	0.0059
	(0.024)	(0.093)	(0.083)	(1774.8)	(0.11)	(0.070)	(0.053)	(0.071)
Under 8 years*Program Area	-0.020	-0.15	0.0039	348.3	0.062	-0.020	0.00057	0.058

	Secondary Progression for 14-17			***Education Expenditure per >5 Child***			***Economic Ladder***	
	% of children gone to sec.	Any child gone to sec.	All children gone to sec.	Ksh spent in previous year	Log of Ksh spent in previous year	Expenditure about median	Overall change	At least 1 step up
	(0.037)	(0.12)	(0.11)	(1671.5)	(0.18)	(0.089)	(0.065)	(0.087)
Wald tests (F)								
Under 8s vs. Over 7s	0.273	0.070	1.020	0.567	0.014	0.160	0.525	0.240
p-value	0.601	0.798	0.312	0.452	0.906	0.685	0.469	0.622
By Baseline Wealth Status (Asset-based)								
Asset Poor*Program Area	-0.026	-0.17*	0.00054	-781.6	0.010	-0.083	0.014	-0.0065
	(0.032)	(0.098)	(0.099)	(1132.1)	(0.15)	(0.079)	(0.058)	(0.076)
Asset Rich*Program Area	-0.043	-0.17	-0.16*	-957.8	0.081	-0.025	0.064	0.060
	(0.027)	(0.11)	(0.094)	(2172.9)	(0.12)	(0.075)	(0.058)	(0.076)
Wald tests (F)								
Asset Rich vs. Asset Poor	0.18	0.00	1.56	0.006	0.147	0.130	0.453	0.440
p-value	0.6706	0.9867	0.2110	0.940	0.702	0.719	0.501	0.509
By Landholding Size								
< 2 acre*Program Area	-0.021	-0.13	-0.0085	-1028.5	0.060	0.035	0.013	0.023
	(0.031)	(0.10)	(0.10)	(1327.0)	(0.13)	(0.079)	(0.055)	(0.080)
≥ 2 acre*Program Area	-0.044	-0.21*	-0.14	-712.9	0.040	-0.14*	0.072	0.039
	(0.029)	(0.11)	(0.096)	(2105.3)	(0.15)	(0.080)	(0.063)	(0.076)
Wald tests (F)								
Under 2 vs. 2 or overs	0.311	0.260	0.860	0.019	0.013	2.760	0.554	0.020
p-value	0.578	0.610	0.353	0.890	0.911	0.097	0.457	0.878
By Dairy Producer								
Dairy*Program Area	-0.028	-0.16	-0.095	-2980.5	-0.048	-0.14*	0.042	0.043
	(0.029)	(0.11)	(0.095)	(2125.5)	(0.14)	(0.081)	(0.065)	(0.084)
No Dairy*Program Area	-0.036	-0.17*	-0.047	1452.4	0.17	0.052	0.036	0.014
	(0.030)	(0.10)	(0.096)	(1400.6)	(0.14)	(0.076)	(0.056)	(0.072)
Wald tests (F)								
Dairy vs. No dairy	0.000	0.010	0.360	3.478	1.351	3.460	0.005	0.080
p-value	0.998	0.942	0.548	0.063	0.246	0.063	0.945	0.780
By Official Position								
Official Position*Program Area	-0.044	-0.23**	-0.11	307.8	-0.013	-0.067	0.061	0.034
	(0.027)	(0.11)	(0.088)	(1981.1)	(0.13)	(0.079)	(0.060)	(0.027)
No Official Position*Program Area	-0.020	-0.10	-0.027	-2019.7	0.099	-0.030	0.013	-0.015
	(0.030)	(0.10)	(0.099)	(1532.6)	(0.15)	(0.076)	(0.056)	(0.028)
Wald tests (F)								
Off. Pos. vs. No Off. Pos.	0.406	0.920	0.510	1.066	0.365	0.140	0.400	1.690
p-value	0.524	0.339	0.476	0.302	0.546	0.705	0.527	0.193

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses and clustered at farmer group level; covariates correlated with programme area (p=<0.1) used in all models; VSZ dummies used for fixed effects in all linear models

Annex 6 : Covariate Comparison of HHs in Programme and Non-programme Areas— Female Participants Only

Characteristic	Program Mean	Non-program Mean	Difference (raw)	Difference (net of county)	Difference (net of VSZ)
Respondent Married	0.64	0.60	0.044*	0.11*	0.12*
			(1.86)	(1.84)	(1.94)
Respondent Widowed	0.20	0.25	-0.042**	-0.14**	-0.14**
			(-2.10)	(-2.05)	(-2.13)
Respondent is Head	0.19	0.24	-0.050**	-0.17**	-0.17**
			(-2.51)	(-2.44)	(-2.44)
Respondent has Official Role	0.46	0.41	0.057**	0.15**	0.15**
			(2.40)	(2.40)	(2.49)
HH reared livestock in 2007	0.68	0.63	0.051**	0.14**	0.14**
			(2.20)	(2.20)	(2.21)
HH member employed in 2007	0.19	0.16	0.035*	0.13*	0.13*
			(1.90)	(1.86)	(1.78)
HH located on tarmac road	0.03	0.05	-0.018*	-0.21*	-0.21*
			(-1.89)	(-1.88)	(-1.88)
Respondent owned main parcel (07)	0.21	0.28	-0.065***	-0.21***	-0.20***
			(-3.12)	(-3.11)	(-3.06)
Highest years of educ. of any adult in HH	10.56	10.85	-0.31*	-0.30*	-0.30*
			(-1.83)	(-1.83)	(-1.83)
Estimated 2007 soil organic carbon (plot avg.)	24.84	23.34	1.49***	1.49***	1.48***
			(3.79)	(4.24)	(5.09)
Elevation (hh level)	1564.95	1522.90	41.7***	42.3***	43.5***
			(3.50)	(3.91)	(4.51)
Observations	**883**	**824**	**1715**	**1715**	**1715**

z/t statistics in parenthesis; VSZ=Village Sampling Zone
* p<0.1, ** p<0.05, *** p<0.01
Probit regression used for net of county and VSZ differences, so coefficients are not directly interpretable, only the *t*-statistics
Only statistics significant differences (p< 0.01) presented from list of 46 covariates