AIR®
AMERICAN INSTITUTES FOR RESEARCH®

Impact Evaluation of 2SCALE: Endline Report

APRIL 2018

Juan Bonilla | Nisha Rai

MAKING RESEARCH **RELEVANT**

AMERICAN INSTITUTES FOR RESEARCH | AIR.ORG

Impact Evaluation of 2SCALE: Endline Report

APRIL 2018

Juan Bonilla | Nisha Rai

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

Copyright © 2018 American Institutes for Research. All rights reserved.

3852_03/18

Contributors

The evaluation of 2SCALE was conducted by American Institutes for Research (AIR). The principal investigators for the overall evaluation were Juan Bonilla and Nisha Rai. The overall team leaders of this report were Juan Bonilla and Nisha Rai; Mitchell Morey and Paul Sirma made important contributions.

The suggested citation for this report is as follows:

Bonilla, J., & Rai, N. (2018). *Impact Evaluation of 2SCALE: Endline Report.* Washington, DC: American Institutes for Research.

Contact information:

Juan Bonilla, PhD
jbonilla@air.org

Nisha Rai, PhD
nrai@air.org

Acknowledgements

We recognize the contributions of many individuals and organisations, without whom it would not have been possible to complete this study. Our thanks go to the International Fertilizer Development Center (IFDC), the Base of Pyramid Innovation Center, the International Centre for development-oriented Research in Agriculture, and Research Solutions Africa (RSA) for the opportunity to carry out this study and for the financial and technical support that they provided.

Our special thanks go to Dr. Jasper Grosskurth (RSA) and Mr. Jasper Gosselt (RSA) for carrying out the data collection and to Mr. Arno Maatman (IFDC) for technical support during the design and fieldwork. The value of the logistical support obtained from the local IFDC coordinators also cannot be overemphasised. Everyone provided valuable logistical support during data collection in the five countries, including programme background information.

Our acknowledgements would be incomplete without mentioning our team of very able enumerators throughout each evaluation country. Specifically, we acknowledge the input of the team of enumerators and supervisors from RSA, whose dedication during data collection ensured that high-quality data were collected. The highly competent team of software programmers at RSA also is gratefully acknowledged.

We also wish to acknowledge the patience exercised by the households, community leaders, and community members from the five countries during interviews. It is our hope that insights from the information they provided will translate into valuable interventions in their communities.

Juan Bonilla, PhD, and Nisha Rai, PhD

Contents

Page

Tables

Page

Figures

Page

1. Introduction, Background, and Objectives

The purpose of this document is to report the results of the final impact evaluation of the Toward Sustainable Clusters in Agribusiness through Learning in Entrepreneurship (2SCALE) programme. The impact estimates use baseline data collected in 2015 and endline data collected in 2017. 2SCALE is an agribusiness project in Africa that aims to connect farmers, buyers, and intermediaries to achieve its goal of improving rural livelihoods and food and nutrition security in Africa. The 2SCALE programme specifically targets smallholder farms in Africa and aims to improve the livelihoods, incomes, and food security of farm households by improving technology, organisational capacity, market access, credit, and extension advice. 2SCALE's market-expanding partnerships aim to provide farmers with incentives to invest in productivity-enhancing technologies. This greater market participation by small-scale local entrepreneurs could then boost food security and agricultural livelihoods in Africa.

The partnership networks that 2SCALE brokers can be classified into two categories:

- Value-chain partnerships are initiated by a non-local-based lead company that operates at international, regional, or national levels and wants to source from (output) or supply to (input) smallholder farmers.
- Agribusiness clusters are grassroots-based partnerships initiated by a local business champion (i.e., farmer group/cooperative, processor, trader, or retailer) that works alongside other relevant stakeholders in the cluster; these stakeholders include smallholder farmers that are involved from the beginning in the partnership negotiations and design.

With funding from the Ministry of Foreign Affairs of The Netherlands, a consortium of organisations, including the International Fertilizer Development Center (IFDC), the Base of Pyramid Innovation Center, and the International Centre for development-oriented Research in Agriculture, launched 2SCALE in June 2012. 2SCALE currently works in nine countries in Africa. In each country, 2SCALE facilitates multiple partnerships to connect farmers, buyers, and intermediaries. The evaluation focused on the impact of one of 2SCALE's partnerships in five of the countries in which it operates: Kenya, Uganda, Ghana, Benin, and Mali. The baseline report also covered Ethiopia, but this country was removed from the final impact evaluation because of external circumstances.[1]

American Institutes for Research (AIR), in partnership with Research Solutions Africa (RSA), conducted a quantitative evaluation of 2SCALE to identify farm-level impacts using a quasi-experimental approach based on matching treatment farmer households to comparison

[1] After baseline data collection, another nongovernmental organisation asked the manager of Solagrow to supply seed potatoes to farmers in the comparison area in Ethiopia. We were planning to proceed with the impact evaluation because the impact evaluation could still provide insights into the impact of the additional components of 2SCALE beyond providing seed potatoes (i.e., the added value of mechanisation services). However, another development occurred that precluded including Ethiopia in the impact evaluation. All of Solagrow's greenhouse and storage facilities were burnt down across the treatment and comparison areas, so 2SCALE postponed activities related to potato production support.

households and tracked them across time. Baseline data and one round of follow-up data (at 24 months) were collected. Baseline data collected in 2015 were analysed as part of the Baseline Report. Endline data were collected in 2017. This Endline Report uses both baseline and endline data. The evaluation was based on the following three research questions (RQs) identified at the beginning of the evaluation:

1. What is the impact of the 2SCALE programme on farmers' income?
2. What is the impact of the 2SCALE programme on food security, including nutritional quality and diet diversity measures?
3. What is the mechanism (e.g., improved technology, organizational capacity, market access, credit, and/or extension advice) through which 2SCALE affects the outcomes of interest?
    a. What is the impact of the 2SCALE programme on commercialization methods and social networks?
    b. Of which component(s) of the 2SCALE programme are farmers aware?
    c. Which component(s) of the 2SCALE programme have farmers used?
    d. Which component(s) of the 2SCALE programme do farmers prefer?

These first two questions relate to the outcomes of interest for the programme. To examine the impact on farmer's income, we consider outcomes that could contribute to increased income. Specifically, we examined investments in crop production (for the target crop specifically and all crops); crop production quantities, revenues, and gross margins (for the target crop specifically and all crops); measures of nonfarm business and credit; and household income sources and, specifically, the past year's noncontract and contract farming income. Although the study design did not allow us to determine the degree to which each programme intervention contributed to the overall impact, the remaining questions helped provide clarity on the impact pathway through which 2SCALE effects change. The main purpose of this Endline Report is to provide insights into the impact of the programme and address the RQs.

The report is organised as follows. Section 2 provides the theory of change, which was previously reported in the Inception and Baseline Reports. Section 3 provides an overview of the study design. Both sections were reported previously in the Baseline Report and are included here for completeness. Section 4 assesses the counterfactual. Section 5 provides answers to the RQs, using all the data collected as part of this evaluation. Section 6 then summarizes the key conclusions.

2. Conceptual Framework: Theory of Change

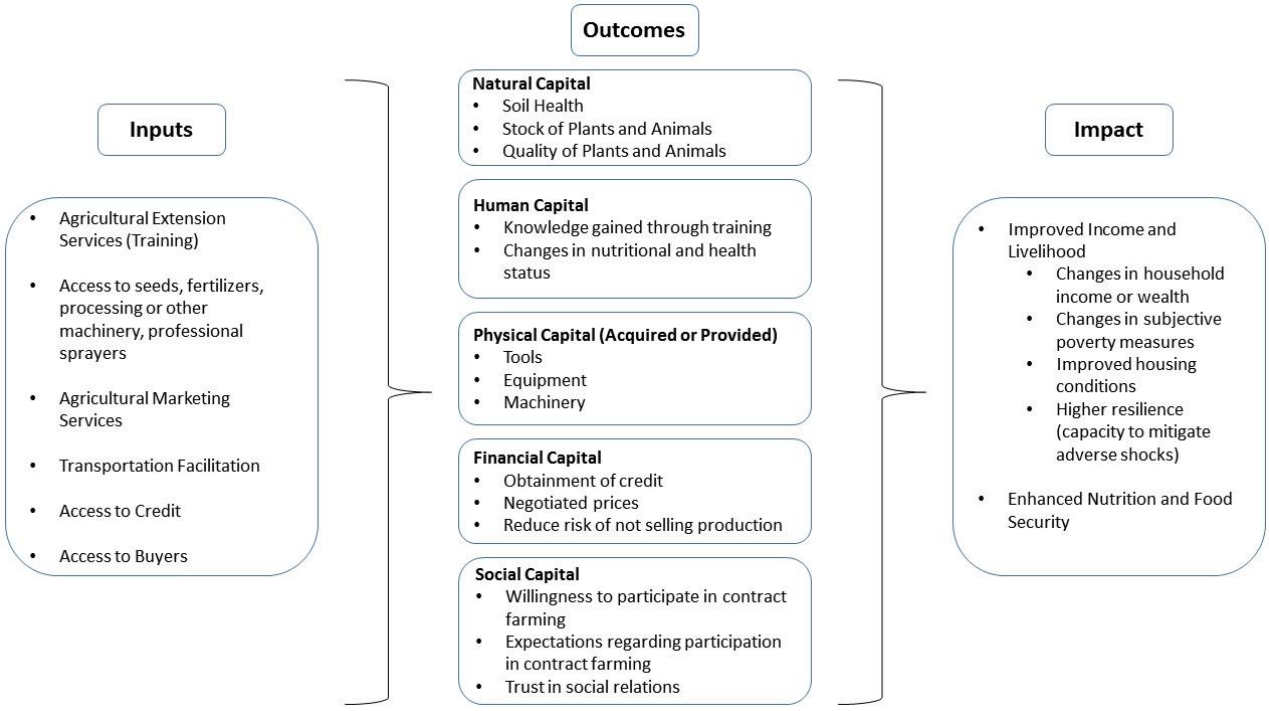
This section presents a simplified theory of change for 2SCALE; a full description of the theory of change is presented as the impact pathway in 2SCALE's strategic work plan for 2015 through 2017 (2SCALE, 2015). The simplified theory of change presents the key elements of 2SCALE's causal logic—how the activities of the programme are intended to bring about improved income and food security outcomes. It is worth noting that 2SCALE operates through public-private partnerships (PPPs), by implementing an inclusive business agenda. Given that partnership-level decisions are being made together with the private partner and very often other stakeholders that participate in the governance of the PPP, each PPP has its own objectives, intervention areas, and dynamics. Thus, strictly speaking, it is not possible to define a unique theory of change of the 2SCALE program for all potential partnerships.

By creating institutional arrangements for coordination, 2SCALE, together with various country stakeholders, identifies the mechanisms that strengthen incentives for co-investment, connectedness, and competitiveness. The arrangements that 2SCALE facilitates constitute a form of value-added processes, which enable the vertical integration of production and distribution systems (Roduner, 2005). In the simplest sense, a value-added process constitutes all activities, institutions, and entities involved in transforming, processing, transporting, and adding value to the product before the product reaches the final consumer (Kidoido & Child, 2014). Along the value-added process, agents exchange the ownership of raw materials, intermediate products, and final products. These different actors also are linked by complex relationships, including demand for goods and services from each other. Under common value-chain models, such as contract farming or farmer cooperatives, farmers obtain access to inputs, credit, technical advice, and market services, which enables them to better integrate into markets and may ultimately result in higher income and food security levels.

The mechanisms or processes that 2SCALE and its partners use to improve agricultural businesses range from the actual provision of inputs (seeds, fertiliser, or machinery) to the provision of services (training, marketing, and transportation) as well as the facilitation of market linkages with sources of credit or potential buyers. Figure 1 is a graphical representation of the general theory of change for 2SCALE. In many cases, 2SCALE and its partners' activities in a specific country are composed of more than one of these mechanisms. These components affect the variety of a farmer's sources of capital. For instance, improved inputs and training could increase a farmer's natural capital by improving the quantity and quality of the land; the training component also could increase the farmer's knowledge and cultural practices. 2SCALE and its partners may increase a farmer's physical capital by facilitating the farmer's access to credit to be able to purchase equipment or directly providing access to machinery. Increased access to buyers resulting from improved market linkages may create the opportunity for predetermined, negotiated prices and reduce a farmer's production risk. Smallholder farmers, by definition, earn income primarily from agricultural production. Should changes in farmers' capital increase agricultural production, then 2SCALE would be associated with improvements in income, assuming the increased linkages facilitate market access for increased production. Improved agricultural outputs and increases in income also should enable farmers to consume a more diverse range of foods, improving food security and nutritional quality.

Built into this theory of change is a set of assumptions governing the pathway from inputs to impact. If the assumptions are unfulfilled, the likelihood of observing the outcomes and impacts decreases. First, we assume no external market disruptions, such as those experienced in the sorghum market in Kenya between 2013 and the first quarter of 2014. Second, we assume that 2SCALE's partners can recruit enough farmers to fully establish market linkages. Lastly, we assume that the lead companies and cooperatives can create sufficient incentives for agents to motivate them to make reasonable arrangements with the farmers, as well as supervise the extent to which the agents uphold the arrangements. Incentivising and supervising the agents is especially important in contributing to farmers' willingness to participate in 2SCALE's programmes.

Figure 1. The Theory of Change



3. Study Design

To answer the RQs, a quasi-experimental approach relying on primary data collection of a farm-level survey was used. The logic of the data collection was generally based on 2SCALE's current field presence and its plans for expansion. In each country, the treatment group was composed of farmers from areas in which 2SCALE was currently operating. 2SCALE farmers were identified through farmer lists. The comparison group was composed of farmers from areas that 2SCALE representatives identified as similar to the treatment areas.

Sample

A multisite longitudinal study was designed to measure the impacts of 2SCALE at the farm level. We conducted a power analysis to determine a sufficient sample size for the study that would enable the detection of meaningful programme effects. This analysis indicated a need for approximately 800 household farms for each country, after accounting for attrition (Bonilla & Rai, 2015). Through the longitudinal design, we surveyed the same farmers in 2017 as in 2015. Multiple reasons motivated the decision to survey the same farmers:

- Surveying the same farmers minimised sample selection issues with the treatment farmers. If the treatment group at endline was based on only existing programme beneficiaries at endline, we would not capture any effects on farmers who were programme beneficiaries at baseline but subsequently dropped out of the programme prior to endline data collection, which would result in inflated impact estimates.
- Surveying the same farmers provided us with values of the outcomes of interest at baseline, which can improve the statistical efficiency through which we calculate the impact.
- Surveying the same farmers provided logistical benefits that aided the data collection at endline by enabling us to provide the data enumerator teams with a list of farmers from baseline.[2] In this way, surveying the same farmers avoided challenges experienced at baseline regarding looking for new farm households.

Because 2SCALE facilitates partnerships at the local level, the sample consists of a diverse portfolio of agriculture product groups, represented in Table 1, where the targeted crop or product group differs by country. 2SCALE's model involves working with partners that are typically either companies or farmer groups.

Table 1. 2SCALE Partnerships for Impact Evaluation

Crop (partnership)	Country	Product group	Partnership type
Sorghum (SHALEM)	Kenya	Staple related	Value chain (output)
Cotton (NYAKATONZI)	Uganda	Cottonseed oil	Value chain (input/output)
Soybeans (processed)	Ghana	Soybean oil seeds	Agribusiness cluster

[2] The farmer list also came with a household information sheet that had key information (GPS coordinates, phone numbers, names, and local names) of the farmers to assist data collection by helping track down farmers.

Crop (partnership)	Country	Product group	Partnership type
Vegetables (EWIT)	Benin	Vegetables/fresh produce	Value chain (input)
Maize (SONAF)	Mali	Staple related	Value chain (output)

Within each country, 2SCALE focused on promoting specific crops and activities.

- **Kenya.** 2SCALE introduced smallholder sorghum farmers to the aggregator Shalem, which buys sorghum on behalf of East African Breweries Limited, to increase efficiency of the sorghum value chain. This partnership also promoted a training innovation in Kenya, the Farmer Field School, in which a group of farmers uses best practices to jointly cultivate a sample plot. The activities also involved development of a financial package and introduction of disease resistant varieties of sorghum.
- **Uganda.** 2SCALE worked with an association of cotton farmers through the Nyakatonzi Growers Cooperative Union. The partnership helped the cooperative procure and install oil milling machinery to produce cottonseed oil products and encouraged farmers to diversify into other oilseeds. The activities also involved agricultural and financial training, developing the market for oil products and secondary products, and increasing access to credit.
- **Ghana.** 2SCALE helped support a variety of activities related to soybean production. The programme linked several producer and processor cooperatives and strengthened these cooperatives through training, increasing access to credit, and setting up a spraying service provider scheme. In addition, for smallholder farmer cooperatives, the programme supported improved production practices, such as access and the use of inputs and equipment, and promoted labor saving equipment and technology. The 2SCALE activities in Ghana also included a pilot marketing programme that marketed soy-based foods to Base of Pyramid Innovation Center consumers to stimulate demand for soy. These activities helped facilitate local entrepreneurs' provision of agricultural inputs and services.
- **Benin.** The lead partner EWIT provides trainings to vegetable farmers on agricultural best practices through learning plots, field visits, training sessions, increased access to finance, and an improved supply of high-quality seeds. The activities also attempted to reduce barriers on access to agricultural inputs, and trained women on entrepreneurship.
- **Mali.** 2SCALE facilitated a partnership with SONAF, a trader of yellow maize that introduced high-yielding maize varieties and provided technical assistance regarding production and quality to smallholder farmers. The activities also attempted to improve post-harvest management and strengthen value chain linkages.

To create the sample at baseline, IFDC's country representatives were asked to identify areas from which approximately 400 treatment and 400 comparison farmers could be enumerated. Either through discussions (in the case of Kenya) or email exchanges we worked with IFDC representatives to select the control area. For each of the countries, we asked key individuals[3] to

[3] We contacted Amos Kisilu (then 2SCALE staff), Ruth Kinoti N'ee Mbogori the partner facilitator in Kenya; Amos Kisilu, Joseph Mwaka (then 2SCALE staff) and Allan Wayira (then 2SCALE staff) in Uganda; Ernest Acheampong and Gabriel Mills in Ghana; Tonato Oliver, C Addupong, Eric Lakoussan, and Ernest Acheampong in Benin; and

fill out a brief questionnaire, designed to give us an idea of the background of the program, where it was currently operating, where it planned to expand, and where it was not working. We wanted to know about areas that were not too near the treatment areas, and we wanted to know about areas where the partnership was not planning to work within the next two years to prevent spillovers (where the program benefits spill over to the comparison group) from biasing the results. However, the areas also needed to have a substantial number of farmers of the target crop to ensure that we would capture an appropriate sample size. Answers to these questions informed our selection of the control area. For further details on the information sources for the comparison areas, see Appendix A.

Within households, adult household members who were mainly responsible for the farming activities in the given household were identified and surveyed. At baseline, enumerators first collected information from a short filter questionnaire to determine if the household should be included. 2SCALE, RSA, and the evaluation team jointly agreed on the criteria for inclusion, as follows. The filter was not necessary at endline because we surveyed the same farmers.

- Confirmation that the household conducts farming activities not exclusively related to livestock. (Comparison farmers had to confirm that they produced the crop for sale outside home consumption; this criterion was necessary only for the comparison farmers because the treatment farmers, through their participation in 2SCALE, already fulfilled this requirement.)
- Have at least 0.25 acre (0.10 hectare) dedicated to the production of the target crop currently and for the most recent harvest. Two considerations determined this threshold. First, the threshold reflected any minimum land size restrictions for participating in 2SCALE that we determined from conversations with 2SCALE's staff in each country. Second, choosing a smaller threshold would make the effect of 2SCALE difficult to detect.
- Have between 0.25 and 10 acres of land for crop production. (Farmers having this many acres could be classified as smallholders, the group 2SCALE expects to benefit from the programme.)
- Relationship with 2SCALE:
  - Treatment farmers were randomly sampled from farmer lists provided by 2SCALE's country representatives.
  - Comparison farmers were sampled using field-based random sampling methods from the identified areas. There was no possibility that farmers in the comparison areas had a relationship with 2SCALE because the farmers were surveyed from areas 2SCALE partners did not target. However, we ensured that although the farmers did not have an agreement with 2SCALE partners, they were willing to have a relationship with the partner if offered.

Frederic Sanogo and Ernest Cheampong in Mali. We also liaised closely with Jan Williem Van Casteren (then 2SCALE staff).

- Willingness to provide cell phone numbers to facilitate finding the farmers at the time of the endline survey.[4]

Table 2. Timeline Information Used to Determine Cut-off Date

Country	Crop	Harvest	2SCALE begin date	Baseline data collection	Reference period baseline harvests	Endline data collection	Reference period endline harvests
Kenya	Sorghum	Feb.–Apr. and Aug.–Oct.	2014	Sept. 2015	Aug. 2014–Aug. 2015	Sept.–Oct. 2017	Aug. 2016–Aug. 2017
Uganda	Cotton	Dec.–Apr.	June 2015	Oct. 2015	Jun. 2014–Jun. 2015	Oct. 2017	Jun. 2016–Jun. 2017
Ghana	Soybeans	Oct.–Nov.	2013[a]	Nov. 2015	Most recent harvest	Nov.–Dec. 2017	Most recent harvest
Benin	Tomatoes and chilies	Tomatoes: Jul., Sept., and Feb. Chilies: Sept. and Jan.	April 2015	Nov. 2015	Most recent harvest[b]	Oct.–Nov. 2017	Most recent harvest
Mali	Maize	Sept.–Oct.	2014	Oct. 2015	Most recent harvest before Aug. 2015	Oct. 2017	Most recent harvest before Aug. 2017

[a]Because no new registrations have been made since 2013, these data do not represent a pure baseline; thus, there was no need to specify a reference period for recent harvests. [b]Because of the different harvest periods for tomatoes and peppers, we asked about the most recent harvest; however, there may be some influence of 2SCALE by the time of the most recent harvest at baseline.

All identified households that did not meet any of the eligibility criteria were substituted with the immediate next household. The final count of households that met the criteria for inclusion in the sample frame at baseline was 4,048: 802 from Kenya, 806 from Uganda, 827 from Ghana, 802 from Mali, and 811 from Benin. These were split essentially equally between the treatment and comparison areas. At endline, the final sample of was 3,840: 756 from Kenya, 748 from Uganda, 776 from Ghana, 800 from Mali, and 760 from Benin. We report on the attrition and relevant attrition analyses in a later section. A limitation with our sample is that we do not know whether buyers would have been willing to cooperate with farmers in the comparison group. If buyers would have been unwilling to cooperate with comparison farmers due to specific characteristics

[4] There was a risk that this restriction might bias our sample from less connected farmers. The percentage of individuals 15 years or older who had a cell phone account in 2014 was 35% in Uganda, 58% in Kenya, 2% in Benin, 13% in Ghana, and 12% in Mali [data from World Development Indicators database (Demirguc-Kunt et al., 2015)]. The mobile account statistic "denotes the percentage of respondents who report personally using a mobile phone to pay bills or to send or receive money through a GSM Association (GSMA) Mobile Money for the Unbanked (MMU) service in the past 12 months; or receiving wages, government transfers, or payments for agricultural products through a mobile phone in the past 12 months." In practice that fraction does not account for individuals sharing cell phone accounts within households, so the penetration fractions are an underestimate at the household or farm level. In addition, RSA did not report challenges with recruiting farmers who met these criteria, suggesting that most of the contacted farmers were able to provide a cell phone number.

of those farmers, then the different characteristics could drive the findings. However, as long as the unobserved characteristics do not vary across time, our empirical specification (described in further detail below) will account for potential self-selection.

Difference in Difference

To conduct a valid assessment of the impact of 2SCALE on farmers' income and food security, it was necessary to establish a clear counterfactual. This required rigorous methodologies that enabled us to address the following question: What would have happened in the absence of the intervention? These methodologies include both randomised controlled trials (RCTs) and quasi-experimental impact evaluations, such as difference in difference (DiD), propensity score matching (PSM), regression discontinuity, and others. By eliminating selection bias and bias from confounding variables, an RCT is the strongest design for making causal claims about programme impacts. However, when an RCT is neither desirable nor feasible, DiD (which compares the average change across time for the treatment group with the average change across time for the comparison group) is a good alternative to an RCT for determining the effectiveness of development programmes. For the evaluation of 2SCALE, the project timeline rendered an RCT infeasible; thus, we used a DiD model for the impact evaluation of 2SCALE. For Ghana, a country for which no pure baseline exists because the programme had been in operation across both treatment and comparison groups since 2013, we used a different identification strategy that combines regression and propensity score matching methods on the endline data. We have included a detailed description of the methodology for Ghana in Appendix B. Although a detailed discussion of our identification strategy was reported elsewhere, particularly in the Inception and Baseline Reports, we include a brief discussion here of the method for completeness.

We performed a DiD analysis to examine the impact of 2SCALE by comparing changes in outcomes across time between programme beneficiaries and nonbeneficiaries. DiD entails calculating the change in an outcome, such as income, between the baseline and the follow-up period for treatment and comparison groups and comparing the magnitude of those changes. In our case, the DiD model used data from before and after 2SCALE's activities to compare the total change in income and food security for households participating in 2SCALE with the total change in income and food security for households in areas in which 2SCALE does not plan to operate. The key assumption underpinning the DiD is that no systematic, unobserved, time-varying difference exists between the treatment and comparison groups. As long as this assumption is satisfied, DiD accounts for potential self-selection from unobserved time-invariant characteristics, which is especially relevant in this evaluation because farmers may self-select into the programme based on characteristics that are not readily available or observable, such as farmers' motivation, ability, or ambition.

Because we did not use randomisation to create the groups, differences will inevitably exist between them, which we discuss in more detail. Having differences between the treatment and comparison groups at baseline is not an issue for a DiD strategy as long at the outcome trends between the two groups are similar over time. That is, as long as the unobserved characteristics do not vary across time, DiD will accounts for potential self-selection. In our impact tables, we

report the impact estimates from a DiD model with no background control variables and a DiD model with control variables, which increases the precision of our estimates.

Quantitative Data Collection

A farmer questionnaire was administered to 4,048 households at baseline and 3,840 households at endline. Appendix C contains RSA's fieldwork report for the endline data collection, which provides details on the data collection process and the difficulties encountered in the field. The relationships that 2SCALE fosters are unique to each country and can thus entail diverse sets of mechanisms and intermediate outcomes. Nonetheless, we structured the survey instrument (included in the Baseline Report) in general terms to maximise applicability across countries. Long-term improvements in farm yields and farm household welfare may result from intermediate behavioural outcomes, such as improved crop husbandry practices and reduced crop damage resulting from changing cultural practices and using inputs (e.g., fertilisers, pesticides) more efficiently. Long-term improvements also could result from improvements in market access and product commercialisation. Farm-level instruments for the impact evaluation were designed to collect both intermediate and long-term outcomes. Seasonal differences in key outcomes, such as income and food security, might exist. By capturing an endline survey 24 months after the baseline, however, the evaluation compared households in the same season as they were surveyed initially, mitigating concerns about seasonal differences.

The intermediate outcomes include variables such as investing in better production inputs, adopting new practices, improving disease/pest management, improving a farmer's power in client negotiations, reducing problems associated with selling the crop, improving access to inputs and activities, and increasing interactions with other farmers or producer group representatives. Beyond these more immediate indicators, the survey included details on all the crops produced by farmers, regardless of land area. Even though 2SCALE focuses on one specific crop in each country, the improved materials or practices that 2SCALE provides may be applicable to other crops the farmer produces or plans to produce. Knowing the total number of crops enabled us to assess whether 2SCALE had an impact on crop production diversity. Along similar lines, the survey included details about food consumption within households to get a sense of whether production diversity or higher income led to more dietary diversity, which was a proxy for food security.

To measure the long-term effects of the programme, the instrument also collected detailed information on crops cultivated in an area larger than 1/10 of an acre (0.04 hectare). These long-term outcomes included crop production amounts and market values, which—along with input expenditures such as fertilisers, pesticides, and labour—enabled us to estimate programme effects on yields, farm productivity, and gross margins.

In the questionnaire, we also investigated farmers' experiences with 2SCALE in terms of the type of relationship, usage, and support from 2SCALE's partner organisations. In the follow-up round of data collection, these items helped us investigate the fidelity with which the programme was implemented in the field.

Finally, we also collected farm household demographic and economic characteristics as well as information on total land area. These and other variables were useful as controls in the analysis and helped improve the match between treatment and comparison households.

Attrition

Attrition within a sample occurs when households originally sampled and interviewed for the Baseline Report were missing or not available for the Endline Report. Mobility, the dissolution of households, death, and divorce can cause attrition and make it difficult to locate a household for follow-up data collection. Attrition causes problems in conducting an evaluation because it not only decreases the sample size (leading to a less precise estimate of programme impact) but also introduces selection bias to the sample, which could lead to incorrect programme impact estimates or change the characteristics of the sample, thus affecting the study's generalisability.

The attrition rates for this study were low (Table 3). Every country's attrition rate was less than 10%, and each country had attrition rates below what we assumed when we conducted power calculations to determine the adequate sample size. Across all the countries, the attrition rate was 5.1%, which translated into 94.9% of the households remaining in the sample from baseline to endline. The main reasons for attrition that RSA reported were migration (which accounted for most of the attrition across the countries), death, and missing/incorrect contact info. In each of the countries, a few people refused the interview because they were either no longer farming or had a disagreement with the aggregator/cooperative. Except for the few cases of interview refusal, based on the reasons for attrition, we would not expect attrition to influence the study results. Nevertheless, we still examined the extent to which the attrition might bias the results. Thus, we conduct attrition analyses for each country.

Table 3. Attrition Rates From Baseline to Endline

	Kenya	Uganda	Ghana	Mali	Benin	Total
Baseline sample	802	806	827	802	811	4,048
Endline sample	756	748	776	800	760	3,840
Difference	46	58	51	2	51	208
Attrition percentage	5.7%	7.2%	6.2%	0.2%	6.3%	5.1%

The two types of attrition are differential and overall. Differential attrition occurs when the treatment and comparison groups differ in the types of individuals that leave the sample. Differential attrition can create biased samples by eliminating the balance between the treatment and comparison groups at baseline. Overall attrition is the total share of observations missing at follow up from the original sample. Overall attrition can change the characteristics of the remaining sample and affect the ability of the study's findings to be generalised to populations outside the study. Ideally, both types of attrition should be small.

Because differences existed between the treatment and comparison groups at baseline, we did not test for differential attrition and report results from an examination of overall attrition only. Examining differential attrition would have involved testing for similarities using baseline data between the treatment and comparison groups for households in both data collections. However,

the Baseline Report indicated that numerous differences existed between the treatment and comparison groups (Bonilla & Rai, 2016). Because an examination of differential attrition is primarily used to determine if attrition eliminated the balance at baseline, such analysis was not relevant for our study, which did not originally exhibit balance.

To examine the extent of overall attrition at endline, we tested similarities using baseline data between the full sample of households at baseline and the sample of households that remained at endline (i.e., the panel sample). Testing these groups on baseline characteristics assessed the extent to which the remaining sample differed from the original sample. In Tables 4–8, we examine the results for control variables and selected outcome.

In general, we did not find significant overall attrition at endline, except in the case of Ghana, where some differences across the samples seemed to exist. As Tables 4–8 show, minimal statistically significant mean differences in the baseline characteristics exist between the remaining sample at the 24-month follow-up and the sample at baseline. In Kenya, we found three of 20 indicators to be statistically different at the 10% significance level. In Uganda, we found one of 16 indicators to be statistically different. In Ghana, we found six of 21 indicators to be statistically different. Benin, we found three out of 13 indicators to be statistically different. In Mali, we found none of the 25 indicators to be statistically different, which is to be expected considering that attrition in Mali was practically non-existent at 0.2% (see Table 3). The statistically significant differences in Kenya and Uganda are results that we could expect resulting purely from chance. These results, and the practically non-existent attrition in Mali, suggest that no bias exists from overall attrition. Although statistically significant differences between the full sample and the panel sample occur in Ghana and Benin, the differences are not practically meaningful in terms of the values. So, overall, across the countries, bias from attrition was not a concern.

Table 4. Overall Attrition Analysis for Selected Variables in Kenya

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
HH size at baseline	5.19	802	5.24	756	0.05***	0.02	0.01
Percentage age between 10 and 14 years	0.15	802	0.16	756	0.00	0.00	0.28
Percentage age between 15 and 64 years	0.56	802	0.55	756	-0.00	0.00	0.17
Male MR	0.45	802	0.45	756	0.00	0.00	0.44
MR's age (years)	43.10	802	43.24	756	0.14	0.11	0.21
How many kilometres from centre of town	14.34	802	14.50	756	0.16	0.10	0.10
Marital status of respondent: married	0.80	802	0.81	756	0.01	0.00	0.20
MR's years farming experience	19.83	802	19.92	756	0.09	0.11	0.40
Language used by respondent: Swahili	0.68	802	0.69	756	0.01*	0.00	0.07

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
Language used by respondent: Meru	0.32	802	0.31	756	-0.01	0.00	0.11
MR is HH head	0.66	802	0.66	756	-0.00	0.00	0.83
MR's education: primary completed	0.36	802	0.36	756	0.01	0.00	0.14
Most important parcel has no erosion	0.40	802	0.39	756	-0.00	0.00	0.40
MR uses cell phone five or more times a day	0.62	802	0.62	756	-0.00	0.00	0.66
Floor: mud/earth only	0.49	802	0.48	756	-0.01	0.00	0.18
Floor: concrete only	0.46	802	0.46	756	0.01	0.00	0.12
Source drinking water: inside compound	0.56	802	0.56	756	-0.00	0.00	0.70
HH main toilet: own pit latrine with slab	0.50	802	0.50	756	-0.00	0.00	0.58
HH main toilet: own pit latrine without slab	0.45	802	0.45	756	0.00	0.00	0.82
Asset index	0.00	802	0.02	756	0.02***	0.01	0.01
Selected outcomes							
Resistant varieties planted	0.75	802	0.76	756	0.01	0.00	0.13
Certified varieties planted	0.80	802	0.80	756	0.01	0.00	0.14
Residue removed prior to planting	0.88	802	0.88	756	-0.00	0.00	0.87
Crops planted early	0.82	802	0.81	756	-0.01	0.00	0.13
Farm suffered from weeds	0.76	802	0.76	756	-0.00	0.00	0.66
Last harvest affected by insects, a fungus, or disease	0.64	802	0.65	756	0.01	0.00	0.12
Percentage quantity affected by pests or disease	0.10	802	0.10	756	0.00	0.00	0.15
Percentage quality or price affected by pests or disease	0.07	802	0.08	756	0.00	0.00	0.23
Log quantity of total harvest	6.15	754	6.16	712	0.01	0.01	0.37
Log value of total harvest	9.40	740	9.41	699	0.01	0.01	0.58
Log gross margins	9.04	644	9.05	608	0.01	0.01	0.57

Note. HH = household; MR = main respondent; SE = standard error.
*$p < .10$. **$p < .05$. ***$p < .01$.

Table 5. Overall Attrition Analysis for Selected Variables in Uganda

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
HH size at baseline	7.66	806	7.66	748	0.00	0.05	1.00
Percentage age between 15 and 64 years	0.46	806	0.46	748	0.00	0.00	0.37
Male MR	0.66	806	0.66	748	-0.00	0.00	0.41
MR's age (years)	42.15	806	42.30	748	0.15	0.15	0.33
How many kilometres from centre of town	2.30	806	2.28	748	-0.01	0.02	0.43
Marital status of respondent: married	0.72	806	0.72	748	0.01	0.00	0.20
MR's years farming experience	21.49	806	21.69	748	0.20	0.14	0.14
Language used by respondent: English	0.32	806	0.32	748	-0.00	0.00	0.50
Most important parcel is owned	0.63	806	0.62	748	-0.00	0.00	0.44
Most important parcel is very fertile	0.50	806	0.49	748	-0.01**	0.00	0.03
Most important parcel has no erosion	0.58	806	0.57	748	-0.00	0.00	0.67
Number of rooms	2.75	806	2.77	748	0.02	0.01	0.11
Walls: pole and mud	0.40	806	0.40	748	0.00	0.00	0.96
HH main toilet: own pit latrine with slab	0.29	806	0.30	748	0.00	0.00	0.36
HH main toilet: own pit latrine without slab	0.52	806	0.52	748	-0.00	0.00	0.61
Asset index	0.00	806	0.00	748	0.00	0.01	0.88
Selected intermediate outcomes							
Resistant varieties planted	0.41	804	0.42	746	0.00	0.00	0.41
Certified varieties planted	0.69	804	0.70	746	0.01	0.00	0.12
Residue removed prior to planting	0.82	804	0.82	746	0.00	0.00	0.91
Crops planted early	0.81	804	0.82	746	0.01	0.00	0.20
Farm suffered from weeds	0.72	804	0.72	746	-0.00	0.00	0.70
Last harvest affected by insects, a fungus, or disease	0.69	804	0.68	746	-0.01	0.00	0.15
Percentage quantity affected by pests or disease	0.14	806	0.14	748	-0.00	0.00	0.21

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
Percentage quality or price affected by pests or disease	0.14	806	0.13	748	-0.00	0.00	0.22
Log quantity of total harvest	6.12	734	6.12	680	0.00	0.01	0.92
Log value of total harvest	13.03	734	13.03	680	0.00	0.01	0.92
Log gross margins	12.80	684	12.80	633	-0.00	0.01	0.86

Note. HH = household; MR = main respondent; SE = standard error.
*p < .10. **p < .05. ***p < .01.

Table 6. Overall Attrition Analysis for Selected Variables in Ghana

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
HH size at baseline	13.77	827	13.74	776	-0.03	0.07	0.64
Percentage age between 15 and 64 years	0.45	827	0.45	776	0.00	0.00	0.43
Male MR	0.57	827	0.58	776	0.00	0.00	0.36
MR is HH head	0.70	827	0.71	776	0.00	0.00	0.77
MR uses cell phone five or more times a day	0.26	827	0.25	776	-0.01*	0.00	0.05
How many kilometres from centre of town	9.31	827	9.09	776	-0.23**	0.10	0.02
MR's years farming experience	19.48	827	19.70	776	0.22**	0.11	0.05
Soil color most important parcel is black/dark	0.72	827	0.72	776	0.00	0.00	0.41
Most important parcel is very fertile	0.74	827	0.73	776	-0.01	0.00	0.12
Number of rooms	5.88	827	5.82	776	-0.06**	0.03	0.04
Roof: grass/straw/thatch/makuti	0.39	827	0.40	776	0.00	0.00	0.36
Roof: metal, tin or zinc	0.41	827	0.41	776	0.00	0.00	0.62
Walls: mud brick	0.68	827	0.68	776	-0.00	0.00	0.66
Floor: mud/earth only	0.67	827	0.66	776	-0.00	0.00	0.75
Floor: mud/earth and concrete	0.25	827	0.24	776	-0.00	0.00	0.65
Drinking water: directly from river/lake	0.25	827	0.25	776	-0.00	0.00	0.67
Drinking water: borehole	0.67	827	0.67	776	-0.00	0.00	0.59
Source drinking water: outside compound	0.76	827	0.77	776	0.01***	0.00	0.01
HH has electricity	0.57	827	0.59	776	0.02***	0.00	0.00

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
HH main toilet: none	0.75	827	0.74	776	-0.01	0.00	0.15
Asset index	-0.00	827	-0.01	776	-0.01	0.01	0.45
Log quantity of total harvest	6.03	812	6.02	762	-0.01	0.01	0.23
Log value of total harvest	6.14	784	6.13	734	-0.01	0.01	0.30
Log gross margins	5.87	730	5.87	682	0.00	0.01	0.98

Note. HH = household; MR = main respondent; SE = standard error.
*p < .10. **p < .05. ***p < .01.

Table 7. Overall Attrition Analysis for Selected Variables in Benin

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
HH size at baseline	7.46	811	7.48	760	0.02	0.04	0.59
Male MR	0.62	811	0.62	760	0.00	0.00	0.63
MR's age (years)	39.94	811	39.91	760	-0.02	0.10	0.81
How many kilometres from centre of town	8.58	811	8.69	760	0.11	0.09	0.22
MR's years farming experience	10.99	811	10.83	760	-0.16*	0.09	0.07
Most important parcel is very fertile	0.63	811	0.63	760	-0.00	0.00	0.56
Most important parcel has no erosion	0.56	811	0.56	760	0.00	0.00	0.65
Number of rooms	2.69	811	2.70	760	0.01	0.01	0.38
Walls: mud brick	0.33	811	0.33	760	0.00	0.00	0.61
Floor: concrete only	0.47	811	0.47	760	0.01	0.00	0.26
HH has electricity	0.38	811	0.39	760	0.01*	0.00	0.10
HH main toilet: own flush in house	0.43	811	0.43	760	-0.00	0.00	0.40
Asset index	0.00	811	-0.02	760	-0.02***	0.01	0.01
Selected intermediate outcomes							
Resistant varieties planted	0.44	811	0.44	760	-0.00	0.00	0.92
Certified varieties planted	0.15	811	0.15	760	0.00	0.00	0.48
Residue removed prior to planting	0.77	811	0.78	760	0.01	0.00	0.17
Crops planted early	0.66	811	0.66	760	-0.00	0.00	0.48
Farm suffered from weeds	0.51	811	0.52	760	0.01	0.00	0.22

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
Last harvest affected by insects, a fungus, or disease	0.19	811	0.19	760	0.00	0.00	0.30
Percentage quantity affected by pests or disease	0.02	811	0.02	760	0.00	0.00	0.83
Percentage quality or price affected by pests or disease	0.02	811	0.02	760	0.00	0.00	0.29
Log quantity of total harvest	6.70	399	6.71	381	0.01	0.02	0.32
Log value of total harvest	12.45	395	12.47	377	0.02	0.03	0.43
Log gross margins	12.54	246	12.56	234	0.02	0.03	0.28

Note. HH = household; MR = main respondent; SE = standard error.
*p < .10. **p < .05. ***p < .01.

Table 8. Overall Attrition Analysis for Selected Variables in Mali

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
HH size at baseline	25.29	802	25.32	800	0.03	0.02	0.17
Percentage age between 15 and 64 years	0.42	802	0.42	800	-0.00	0.00	0.78
Male MR	0.95	802	0.95	800	-0.00	0.00	0.17
MR's age (years)	47.34	801	47.34	799	-0.00	0.02	0.90
How many kilometres from centre of town	35.36	802	35.38	800	0.01	0.01	0.17
Marital status of respondent: married	0.99	802	0.99	800	-0.00	0.00	0.19
MR's years farming experience	24.25	802	24.27	800	0.02	0.02	0.23
Language used by respondent: Bambara	0.64	802	0.64	800	0.00	0.00	0.16
MR is HH head	0.92	802	0.92	800	-0.00	0.00	0.16
Most important parcel is very fertile	0.37	802	0.37	800	0.00	0.00	0.16
Most important parcel has no erosion	0.57	802	0.57	800	0.00	0.00	0.84
Number of rooms	5.51	802	5.52	800	0.01	0.01	0.16
Shelter type: traditional house	0.83	802	0.83	800	-0.00	0.00	0.16
Roof: grass/straw/thatch/makuti	0.39	802	0.38	800	-0.00	0.00	0.16
Roof: iron sheets	0.54	802	0.54	800	0.00	0.00	0.16
Walls: pan brick	0.53	802	0.53	800	-0.00	0.00	0.16
Walls: mud brick	0.35	802	0.35	800	0.00	0.00	0.16

	Full sample		Panel sample		Balance test		
Dependent variable	**Mean**	**N1**	**Mean**	**N2**	**Diff**	**SE**	***p*-value**
Floor: mud/earth only	0.77	802	0.77	800	-0.00	0.00	0.16
Drinking water: protected well	0.41	802	0.41	800	-0.00	0.00	0.16
Source drinking water: inside compound	0.29	802	0.29	800	0.00	0.00	0.16
Source drinking water: outside compound	0.49	802	0.49	800	-0.00	0.00	0.16
HH has electricity	0.19	802	0.19	800	0.00	0.00	0.16
HH main toilet: own flush in house	0.70	802	0.70	800	-0.00	0.00	0.16
HH main toilet: own flush outside house	0.28	802	0.28	800	0.00	0.00	0.16
Asset index	-0.00	802	0.00	800	0.00	0.00	0.39
Selected intermediate outcomes							
Resistant varieties planted	0.30	798	0.30	796	0.00	0.00	0.16
Certified varieties planted	0.60	798	0.60	796	0.00	0.00	0.78
Residue removed prior to planting	0.54	798	0.54	796	0.00	0.00	0.91
Crops planted early	0.45	798	0.45	796	-0.00	0.00	0.88
Farm suffered from weeds	0.23	798	0.22	796	-0.00	0.00	0.50
Last harvest affected by insects, a fungus, or disease	0.11	798	0.11	796	0.00	0.00	0.16
Percentage quantity affected by pests or disease	0.01	802	0.01	800	0.00	0.00	0.17
Percentage quality or price affected by pests or disease	0.01	802	0.01	800	0.00	0.00	0.17
Log Quantity of total harvest	8.13	744	8.13	742	-0.00	0.00	0.17
Log value of total harvest	13.15	578	13.15	576	-0.00	0.00	0.21
Log gross margins	12.89	540	12.89	538	-0.00	0.00	0.16

Note. HH = household; MR = main respondent; SE = standard error.

4. Assessing the Counterfactual

The purpose of using rigorous quantitative methodologies to assess impact is to ensure that if the approach is successful, it creates a reasonable counterfactual where the treatment and comparison groups are similar in terms of their observable and unobservable characteristics. The only remaining observable difference, then, is that the treatment group received the programme, enabling us to assess the programme's impact. It could be that farmers self-select into the programme based on characteristics that are not readily available or observable, such as farmers' motivation, ability, or ambition, then our estimate of the effect would be biased. By using DiD, however, we mitigated concerns about this selection bias because DiD accounts for any potential self-selection if those unobserved characteristics do not vary across time.

In the Baseline Report, we examined the extent to which our design created a reasonable counterfactual. In this report, we also present information relevant to assessing the counterfactual with the exception that we did not include the balance checks reported at baseline, where we examined the mean differences between the treatment and comparison groups across variables prior to matching. Balance checks are especially important to examine at baseline because the baseline values represent starting points prior to the programme. In this section, we first describe the outcome and control variables used in the analysis. Second, we summarize the balance checks between the treatment and comparison groups we performed at baseline. And third, we provide and discuss maps of the treatment and comparison areas.

Construction of Variables

Outcome Variables

Table 9 presents the most relevant intermediate and final outcomes. It describes the type of variable—whether categorical or continuous and the units of measurement—as well as the level at which the variable can be constructed. For the analysis of baseline data, outcome variables were created at three different levels. The first level was the crop level; the second level resulted after aggregating crop data by contracted status, namely, targeted or nontargeted; and the third level was when all crops were aggregated at the household level. These different levels enabled us to better characterise the agricultural conditions of households with multiple crops. Note, for instance, that the value of a specific variable (say, gross margins) varies by level only if the household has more than one crop. Otherwise, having only one crop yields the same value at all levels because the crop level is the same as the household level.

Table 9. Description of Selected Outcome Variables[a]

Intermediate outcomes	Variable type	Variable level
Cultural practices		
Crop rotation, early planting, intercropping, removal of plant residue, planting resistant varieties, use of certified planting material, crop monitoring, and weeding	Yes = 1, No = 0	H
Transactional		
Farmer having negotiation power in client transactions, reporting of crop selling problems, and reporting of accessing inputs and activities	Yes = 1, No = 0	H
Inputs		
Value of seed planted (imputed)[b]	Local currency	C; T, NT; H
Organic fertiliser used	Yes = 1, No = 0	C; T, NT; H
Inorganic fertiliser used	Yes = 1, No = 0	C; T, NT; H
Value of inorganic fertiliser used	Local currency	C; T, NT; H
Pesticide used	Yes = 1, No = 0	C; T, NT; H
Value of pesticides used	Local currency	C; T, NT; H
Biological crop protection used	Yes = 1, No = 0	C; T, NT; H
Value of biocontrol used	Local currency	C; T, NT; H
Total family labour days	No. days	C; T, NT; H
Total paid labour days	Local currency	C; T, NT; H
Value of paid labour	Local currency	C; T, NT; H
Final outcomes		
Household well-being		
Past year's programme and nonprogramme farm income, agricultural, nonagricultural, and other income wage earnings	Local currency	H
Subjective poverty measures	Various scales	H
FANTA dietary diversity and food security	Various scales	H
Yields and productivity		
Quantity	Weight units	C; T, NT; H
Total value of harvest	Local currency	C; T, NT; H
Total value of harvest consumed (imputed)	Local currency	C; T, NT; H
Total value of harvest sold	Local currency	C; T, NT; H
Gross margins	Local currency (see formula)	C; T, NT; H
Production diversity		
Number of crops produced	Count	T, NT; H
Shannon index	See formula	H
Simpson index	See formula	H

Note. C = crop; FANTA = Food and Nutrition Technical Assistance Project; H = household level; NT = nontargeted crop; T = targeted crop.

[a]For numeric outcome variables (i.e., those that do not take yes or no values), we set as missing those values above the 99th percentile and those below the 1st percentile to control for outliers. [b] In general, we calculated the imputed value of seed planted using the median price of purchased seed. At endline in Mali, only 86 farmers reported purchasing seed for improved or local maize. Of these farmers, only three farmers of improved maize (and none of the farmers of local maize) reported costs they spent on purchased maize. If we had used the endline values, the value of seed planted would have been 0 because so few farmers who purchased seeds reported costs for these purchases. Instead, for the imputed value of improved and local maize at baseline and endline, we used the average price of purchased seed at baseline. It was necessary to use the average price for maize because of significant outliers.

As shown in Table 9, some outcome variables, such as inputs used (e.g., seeds, fertilisers, pesticides, and labour), yields, and productivity (e.g., value of harvest and gross margins) could be analysed at the crop level. In addition, most outcomes could be aggregated into either targeted or nontargeted crop and then totalled by household. The most notable exceptions are the two indices on production diversity, which are more meaningful at the household level.

Intermediate Outcomes

Because long-term improvements in farm yields and farm household welfare may result from short-term outcomes, such as improved crop husbandry practices, market access, and product commercialisation, we captured potential intermediate outcomes of 2SCALE. The most relevant intermediate outcomes are indicator variables for cultural practices used, the use and value of inputs, indicators for farmer's negotiation power, reporting of crop selling problems, and reporting of accessing inputs and activities.

Final Outcomes

Beyond these intermediate indicators, the survey listed all the crops produced by farmers, regardless of land area. This enabled us to assess whether the 2SCALE programme had an impact on crop production diversity. Along similar lines, the survey included details about food consumption within households to get a sense of whether production diversity or higher income led to more dietary diversity, which was a proxy for food security. One such measure of food security was developed by the Food and Nutrition Technical Assistance Project (FANTA).

Specifically, we used a method implemented by FANTA based on the idea that being food insecure causes predictable responses that can be captured in a survey and summarised in a scale (Coates, Swindale, & Bilinsky, 2007). Some of the reactions explored were related to feelings of uncertainty or anxiety over food, perceptions that the quantity of food available is insufficient, perceptions that food was of insufficient quality, and reductions of food intake for adults and children. For example, a question relating to perceptions of insufficient quantity asked whether any adults had to eat less than they thought they should. FANTA identified a set of questions that appeared to distinguish food secure from food insecure households across different cultural settings around the world. Each question was asked with a recall period of 4 weeks, where the respondent was first asked whether the condition in the question happened at all. If the respondent answered "yes," a frequency-of-occurrence question was asked to determine whether the condition happened rarely (once or twice), sometimes (three to 10 times), or often (more than 10 times) in the past 4 weeks. Responses to these questions were summarised in a scale and provide a continuous indicator of the degree of a household's food insecurity. Greater values of the food security measure indicate more food insecurity.

FANTA also includes guidelines for a dietary diversity questionnaire that can be used at the household level. Specifically, this diversity measure involves calculating dietary diversity scores by summing the number of food groups consumed by anyone in the household during a reference period (our questionnaire used the last 7 days). Thus, the dietary diversity scores consisted of a simple count of food groups that the household consumed in the last 7 days.

To measure the long-term effects of the programme, the instrument also collected detailed information on crops cultivated in an area larger than 1/10 acre (or 0.04 hectare). These long-term outcomes included crop production amounts and market values, which—along with input expenditures such as fertilisers, pesticides, and labour—allow us to estimate programme effects on yields (i.e., quantities); farm productivity (i.e., quantities per unit of cultivated area); and gross margins (*G*), which for crop *c* in household *h* is defined as follows:

$$G_{ch} = REV_{ch} - L_{ch} - B_{ch} - P_{ch} - F_{ch} - S_{ch}$$

where REV refers to the revenues from the quantity harvested. From these revenues, we subtracted the costs associated with hired labour (L), the biological control[5] used (B), the pesticide used (P), any organic and inorganic fertiliser used (F), and seed planted (S).

All values collected from the farmer survey that asked about the total value per crop (e.g., total cost of pesticides, total harvest revenue received). Following common practice, missing prices, on the other hand, were imputed (as necessary) using the median reported price for a given commodity or crop at the local level. If too few observations were available for a specific crop and input at the local level, imputations were done using the median prices at the next available aggregation level (i.e., the national level). In addition, all area variables were converted to a standardised unit (acres for all countries) to express all monetary values per unit of area to facilitate the comparison of farm households with different land extensions.

Questions about nonfarm businesses and credit were used to provide a picture of the financial position of the households; changes in access to credit or financial inclusion might occur because of 2SCALE. In addition, respondents' subjective poverty was an outcome variable being considered. Subjective poverty measures can be constructed by asking respondents their perception of consumption or their income classification. Research has shown that poverty lines constructed from subjective measures similarly match those based on more objective poverty calculations (Pradhan & Ravallion, 2000).

Lastly, we created two different production diversity indices: the Shannon and Simpson indices. Crop counts often are criticised as measures of production diversity because they fail to capture the area planted and therefore the relative importance of different crops. These indices are commonly used to measure production diversity and incorporate, in distinct ways, land area planted with different crops (Smale, 2006). The Shannon index captures the evenness and proportional abundance of crops and is defined as follows:

$$\text{Shannon} = -\sum \alpha_i \ln \alpha_i$$

where α_i is the share of land area planted with a particular crop. This measure has a lower limit of zero, which occurs when only one crop is planted. The Simpson index also captures the relative abundance of each crop and is defined as follows:

$$\text{Simpson} = 1 - \sum \alpha_i^2$$

[5] Biological control agents include macrobials (good insects), microbials (good microorganisms), associated products (e.g., traps), or biochemicals (plant extracts).

which also has a lower limit of zero, and where α_i is defined as the share of land area planted with a specific crop. Along with the crop count, the Shannon and Simpson indices provide estimates of diversity, with higher values suggesting greater production diversity.

Control Variables

In the questionnaire, we collected a rich set of variables that enabled us to characterise households in the sample and match treatment households to comparison households. In addition, these variables were used as control variables in the analyses. Further, the control variables enabled a more precise estimate of programme impacts because they are good predictors of the intermediate and final outcomes of interest.

In all five countries, variables at baseline that we used as control variables included household size, the main respondent's years of farming experience, the kilometres away from the centre of town, an indicator that main respondent is male, and an asset index. The full set of variables differed by country based on the variation in responses of that variable in each country and the resulting explanatory power that the individual variables had on explaining treatment status. In general, the other variables covered demographics, housing, water and sanitation, and agriculture characteristics.

Balance Tests: Treatment and Comparison

In the Baseline Report, we examined the baseline quantitative data to determine the extent to which the treatment and comparison groups were similar and understand the main differences. Had randomisation been possible, we would have expected a balance of outcome and control indicators between the two conditions (i.e., treatment and comparison). That is, we would have expected the mean outcomes between both groups to be similar. With nonexperimental methods, such as DiD, however, it is unlikely that the treatment and comparison groups will be balanced in their mean outcome, and, indeed, the balance tests reported at baseline revealed that differences between treatment and comparison farmers existed along various dimensions shown to be predictive of programme participation (Bonilla and Rai, N., 2016). Using *t*-tests of differences in means across groups, at baseline we tested all the outcome measures and control variables for statistical differences between the two groups, and we found that for most outcomes and control variables the average characteristics of treatment and comparison groups were statistically different. Given the quasi-experimental design of the study, these differences were expected. Furthermore, having differences between the treatment and comparison groups at baseline is not an issue for a DiD strategy as long at the outcome trends between the two groups are similar over time. That is, the DiD strategy does not rely on the characteristics being the same at baseline and instead requires that both groups follow the same trend over time. Nevertheless, these differences in observable characteristics emphasize the relevance of controlling by those characteristics in our specifications.

Maps of Treatment and Comparison Areas

The maps included in Figures 2-6 plot the locations of the treatment and comparison farmers for each of the five countries. All of the maps include a zoomed-out depiction of the country with

the relevant area highlighted and a zoomed-in version of the actual treatment and comparison farmer locations. In Kenya, Benin, and Mali the distances between treatment and comparison farmers are relatively small. However, those distances are larger in Uganda and Ghana, suggesting that some differences between the farmers may exist. Additionally, we understand that the comparison group in Uganda included farmers that also focus on other crops and that some farmers have subscriptions to the Masindi District Farmer Association. While these facts suggest that some differences existed between farmers in the two groups, as long as both groups follow the same trend in outcomes over time, the DiD strategy will still serve to estimate the impact of 2SCALE for Uganda. The separation between the treatment and comparison areas suggests that the risk of spillovers biasing the results should be minimal.

Figure 2. Map of Treatment and Comparison Areas (Kenya)

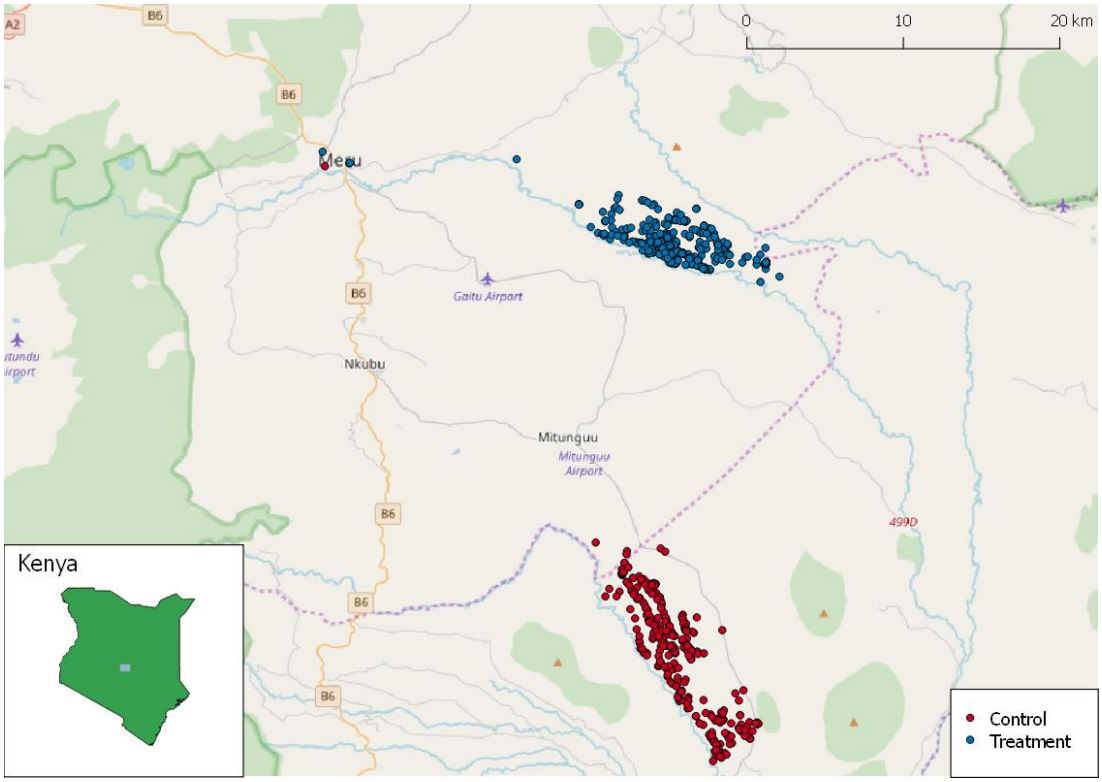


Figure 3. Map of Treatment and Comparison Areas (Uganda)

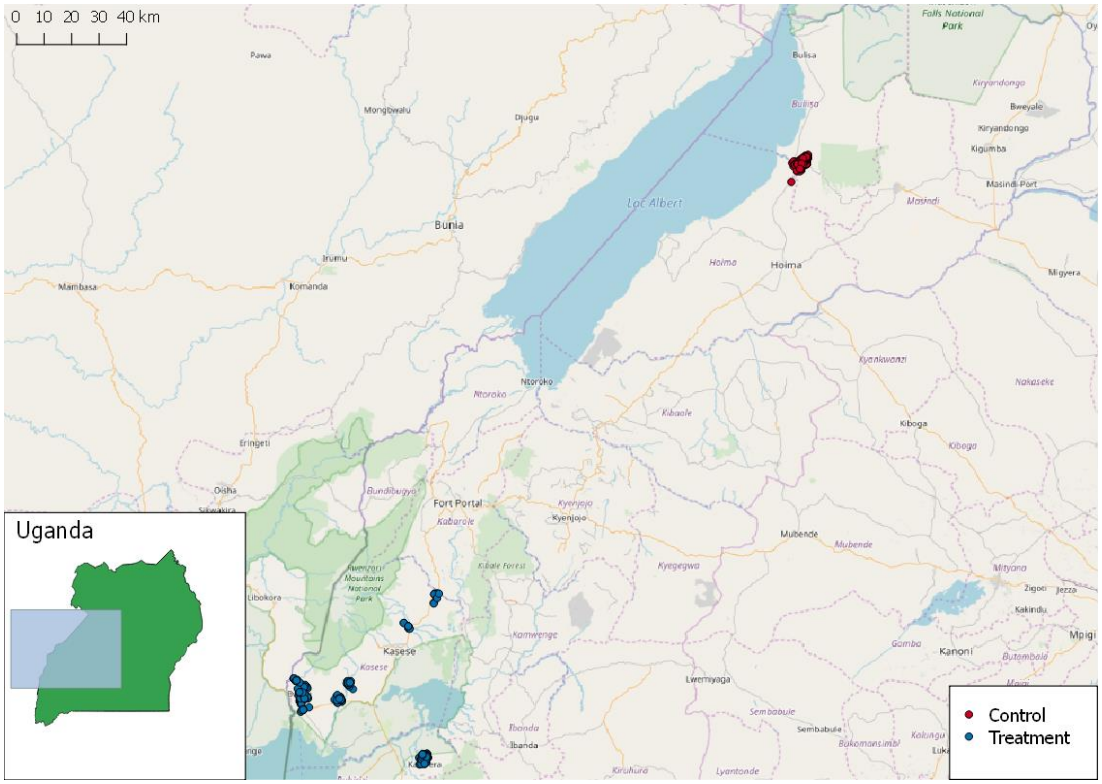


Figure 4. Map of Treatment and Comparison Areas (Ghana)

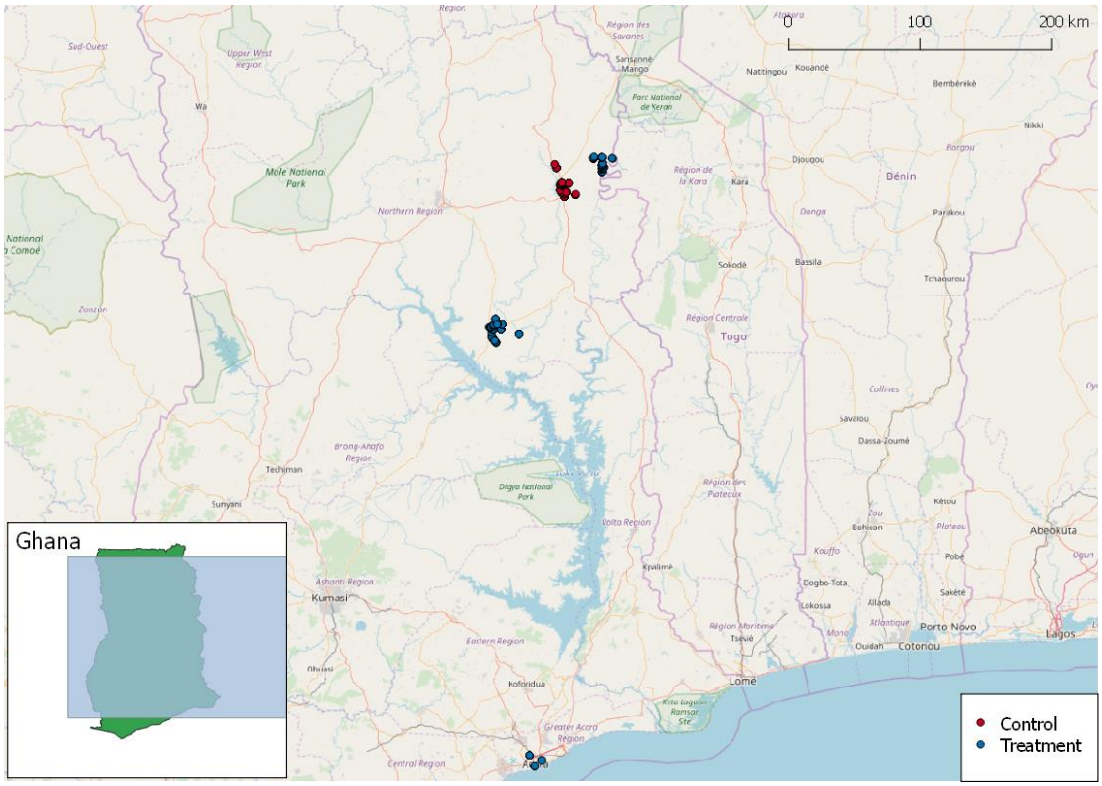


Figure 5. Map of Treatment and Comparison Areas (Benin)

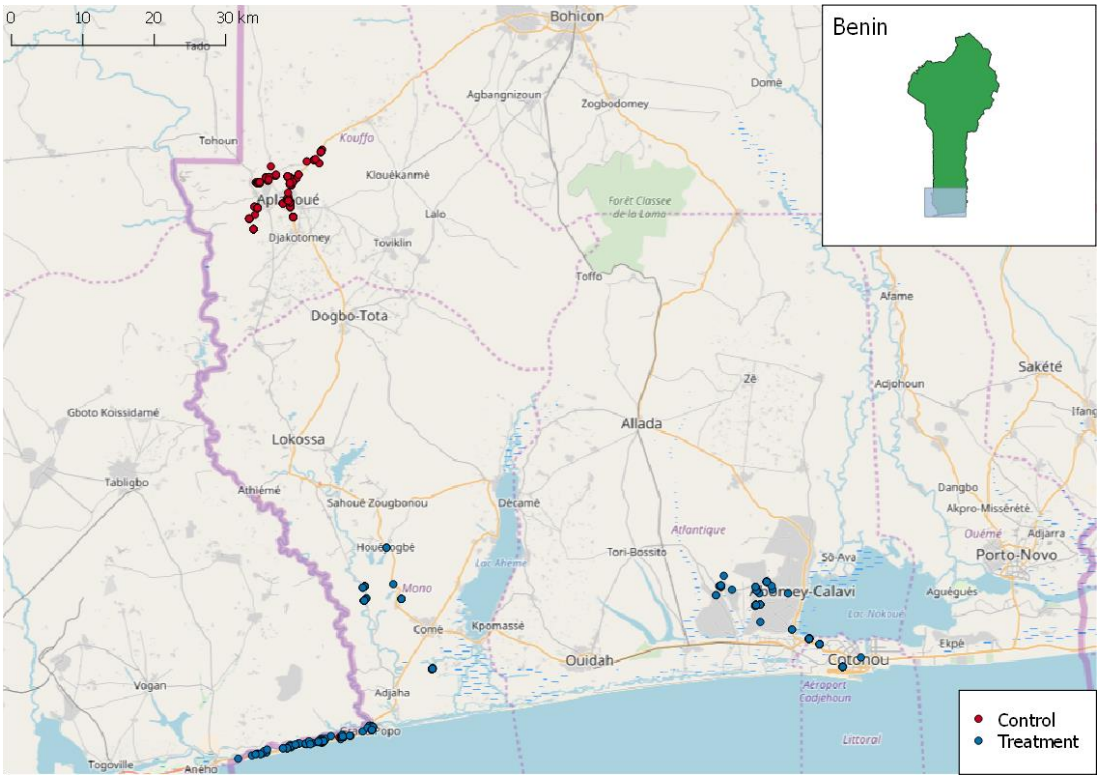
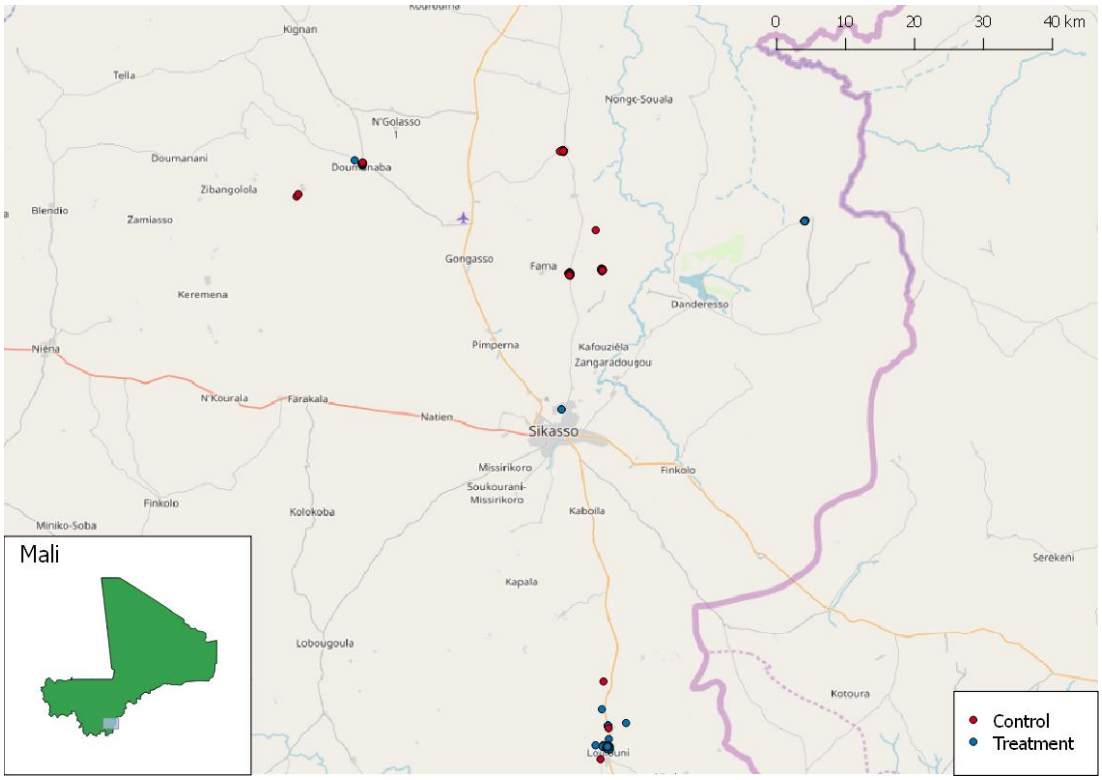


Figure 6. Map of Treatment and Comparison Areas (Mali)



5. Endline Insights into the Research Questions

Identifying the impact of 2SCALE requires careful analysis of both baseline and endline data. The information collected at 24 months provides insights into the RQs that particularly related to the impact of 2SCALE across income and food security. However, the data also provided some clarity on the mechanism through which 2SCALE affects the outcomes of interest. We organise this section by discussing our overall approach to the research questions, presenting the results for each country, and examining the cross-country descriptive data.

We have organised the impact tables such that the first column (No Covs) presents the DiD estimates from a specification that includes no covariates. The estimates in the second column (Covs) are from a DiD specification that includes control variables. By including control variables that have explanatory power in explaining variation in the outcome variables, the precision with which we can estimate the impact improves with this second specification relative to the first.

We examined the impact estimates for consistency across the two specifications in terms of the statistical significance of the impact estimate as well as the value of the estimate. In reporting the statistically significant impacts, we focused on reporting those from the DiD specification with covariates. DiD with covariates compared groups that we knew had statistically significant differences at baseline; however, by controlling for additional covariates this specification helps account for these differences in observable characteristics between the treatment and control groups.

RQ1: What is the impact of the 2SCALE programme on farmers' income?

To examine the impact of 2SCALE on farmer's income, we looked at a variety of crop and household outcomes. However, the exact set of outcomes that we examined varied by country as we tailored the outcomes to the specific activities the partnerships provided. In general, we first looked at intermediate outcomes that could contribute to an increase in income. Specifically, we examined intermediate outcomes related to changes in cultural practices if the partnership involved training that might have influenced these practices. Second, we looked at outcomes related to the target crop in each country, such as the gross margins per unit of area, the quantity of the total harvest per unit of area, and the value of the total harvest per unit of area, which are measures of productivity, in addition to the inputs that determine gross margins if the partnership focused on increasing access to inputs. Third, because the benefits that 2SCALE provides toward production of the target crop may spill over into the production of other crops, we also examined crop diversity measures. Finally, additional outcomes related to farmer income include those related to nonfarm business and credit and household income sources. It is worth noting that borrowing debt is neither unambiguously good or bad. An increase in borrowing debt could relate to farming expansion, whereas a decrease in borrowing debt could suggest that the farmer is not capital constrained. However, we would expect to see an increase in borrowing or debt for partnerships that increased access to credit or finance.

For the variables cost of seed planted, cost of pesticide, cost of paid labour, cost of paid plus family labour, the quantity of the total harvest (kilograms), the value of the total harvest, and gross margins per unit of area, we present the estimates based on natural logarithm transformed outcome variables, where the transformed outcome variable equals the natural log of the original outcome variable. While we controlled for outliers during our analysis, the natural log transformation helped further account for any outliers present with these variables. With this transformation, the interpretation of the impact of the programme was that the outcome changed by 100*(impact estimate) percent, all else being equal.

RQ2: What is the impact of the 2SCALE programme on food security, including nutritional quality and diet diversity measures?

To examine the impact of 2SCALE on food security, we consider both subjective measures and actual reports of the specific food items consumed. For each country, we report the impact estimates of subjective measures of poverty and food security as well as the food insecurity scale index, where greater values of the food insecurity measure indicate more food insecurity. Additionally, we report the impact estimates of specific food items consumed and the resulting FANTA household dietary diversity score.

In the following table, we present the main intermediate and final outcomes relevant for each country, the expected impact based on the programme's log frame, as well as the main reason why we include a specific outcome for the analysis, which is based on the partnership activities. As noted before, 2SCALE operates through public-private partnerships (PPPs), each one with their own objectives, intervention areas, and dynamics. Thus, it can be argued that each PPP has its own theory of change operating though some general programme guidelines and goals. The following table is an attempt to summarize the relevant outcomes for each PPP.

Table 10. Country-specific outcomes of interest and expected impact

Outcomes	Expected Impact	Relevant for	Impact due to programme's
Parcels and Cultural Practices			
Resistant varieties planted	(+)	K, U, B, M	Focus on quality/disease resistance
Certified varieties planted	(+)	K, U, B, M	Focus on quality/disease resistance
Residue removed prior to planting	(+)	K, U, B, M	Trainings received
Crops planted early	(+)	K, U, B, M	Trainings received
Farm suffered from weeds	A	K, U, B, M	Trainings received
Last harvest affected by insects, a fungus, or disease	(-)	K, U, B, M	Focus on quality/disease resistance
Percentage quantity affected by pests or disease	(-)	K, U, B, M	Focus on quality/disease resistance
Percentage quality or price affected by pests or disease	(-)	K, U, B, M	Focus on quality/disease resistance
Nonfarm Business and Credit			
Borrowed on credit from someone outside the household	A	K, U, B, G	Financial package received

Debt from loans contracted in the past 12 months	(+)	K, U, B, G	Financial package received
Tried to borrow from someone outside the household	(-)	K, U, B,	Financial package received
Would apply for a loan if certain he will get it	(-)	G	Financial package received
Product commercialisation			
Number of clients sold to		B, M	
Client facilitaded access to agricultural training	(+)	K, B, M	Trainings received from client
Client facilitaded access to credit	(+)	K, B,	Financial package received from client
Client facilitaded access to fertiliser	(+)	B	
Client facilitaded access to seeds	(+)	B	Financial package received from client
Networks and Social Capital			
Group facilitated access to seeds	(+)	B, G	Promotion of input access
Group facilitated access to fertiliser	(+)	B, G	Promotion of input access
Group facilitated access to machinery	(+)	U, B, G	Promotion of input access
Group facilitated access to professional sprayers	(+)	B, G	Set up of spraying service scheme
Group facilitated access to labour	(+)	B, G	Promotion of input access
Group facilitated access to credit	(+)	U, B, G	Promotion of financial access
Group facilitated access to agricultural training	(+)	U, B, M	Focus on training
Group facilitated access to marketing	(+)	U, B, M, G	Assistance with marketing strategy
Group facilitated access to processing	(+)	U, M	Assistance with output processing
Group facilitated access to distribution	(+)	B, M	Support for value chain coordination and networking
Final outcomes			
Quantity of the total harvest (kilograms)	(+)	All	Emphasis on longer-run outcomes with time
Value of the total harvest	(+)	All	Emphasis on longer-run outcomes with time
Gross margins	(+)	All	Emphasis on longer-run outcomes with time
Producing target crop	(+)	All	Emphasis on longer-run outcomes with time
No. of crops produced	A	All	Emphasis on longer-run outcomes with time
Food security	(+)	All	Emphasis on longer-run outcomes with time
Dietary diversity score	(+)	All	Emphasis on longer-run outcomes with time

Notes: K=Kenya; U=Uganda; B=Benin; M=Mali; G=Ghana. (+) Means that the expected programme impact is to increase the value of the outcome. (-) Means that the expected programme impact is to decrease the value of the outcome. A=Ambiguous

RQ3: What is the mechanism (e.g., improved technology, organisational capacity, market access, credit, and/or extension advice) through which 2SCALE affects the outcomes of interest?

Although the study design does not allow us to determine the degree to which each programme intervention contributes to the overall impact, we examine the mechanism through which 2SCALE affects outcomes by looking at the effect of the programme on intermediate outcomes. Specifically, we focus on the impact of 2SCALE on non-independent product commercialisation in addition to networks and social capital as applicable in each country/partnership. Examining these additional outcomes can provide insight into the channels through which 2SCALE operates. In a separate section, we examine the extent to which treatment farmers are aware of and use the programme and the preferences they have for programme components. Our examination of awareness, use, and preferences is descriptive in nature only and cannot be used to infer causal relationships.

Kenya

In Kenya the partnership with Shalem aimed to increase efficiency of the sorghum value chain. The activities included training (through the farmer field school), developing a financial package, and introducing disease resistant varieties of sorghum. We examine the impact of the programme on cultural practices, which may have been influenced by the training; sorghum production and crop diversity; access to credit, which may have been influenced by the financial package; income from the contracted crop; food security measures; and non-independent product commercialization since the 2SCALE model involved client-relationships.

Parcels and Cultural Practices (Intermediate Outcomes)

Table 11 reports the impact estimates for using a variety of cultural practices in Kenya. Overall, we found statistically significant increases in the use of cultural practices by treatment farmers, which could be linked to the training received from the farmer field schools. Specifically, treatments farmers were 21 percentage points more likely to plant resistant varieties. Weak evidence (at the 10% significance level) showed an 8-percentage point increase in planting with certified varieties and a 2-percentage point reduction in the quality or price of their crops being affected by pests or disease. There was no statistically significant difference in the likelihood of removing residue prior to planting. At the 10% significance level, treatment farmers were 7 percentage points more likely to plant their crops early.

Table 11. Parcels and Cultural Practices (Kenya)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Resistant varieties planted	0.21***	0.21***	0.83	0.68	0.83	0.47	1512

Dependent variable	DiD estimates No Covs (1)	DiD estimates Covs (2)	Baseline means T (4)	Baseline means C (5)	Endline means T (6)	Endline means C (7)	*N*
	(4.62)	(4.69)					
Certified varieties planted	0.08*	0.08*	0.96	0.64	0.90	0.51	1512
	(1.91)	(1.94)					
Residue removed prior to planting	0.04	0.04	0.94	0.83	0.92	0.77	1512
	(1.06)	(1.09)					
Crops planted early	0.07*	0.07*	0.81	0.82	0.85	0.79	1512
	(1.80)	(1.92)					
Farm suffered from weeds	0.01	0.01	0.77	0.74	0.81	0.76	1512
	(0.19)	(0.23)					
Last harvest affected by insects, a fungus, or disease	0.00	0.00	0.66	0.64	0.74	0.72	1512
	(0.06)	(0.08)					
Percentage quantity affected by pests or disease	0.00	0.00	0.09	0.11	0.12	0.13	1512
	(0.20)	(0.22)					
Percentage quality or price affected by pests or disease	-0.02*	-0.02*	0.06	0.09	0.09	0.13	1512
	(-1.68)	(-1.66)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Target Crop Outcomes

In Kenya, statistically significant evidence indicates that farmers' sorghum production increased (Table 12). 2SCALE farmers had statistically significant (at the 1% level) increases in the quantity of the total harvest (90%), the value of the total harvest (89%), and the gross margin (85%).

Table 12. Impact Estimates Target Crop (Kenya)

Dependent variable	DiD estimates No Covs (1)	DiD estimates Covs (2)	Baseline means T (4)	Baseline means C (5)	Endline means T (6)	Endline means C (7)	*N*
Quantity of the total harvest (kilograms)	0.91***	0.90***	6.46	5.34	6.64	4.61	1276
	(8.23)	(8.23)					
Value of the total harvest	0.90***	0.89***	9.71	8.57	10.03	8.00	1259
	(8.11)	(8.09)					
Gross margins	0.81***	0.85***	9.35	8.17	9.74	7.74	1062
	(5.42)	(5.65)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. Only panel observations are included. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

All Crop Outcomes

As we would expect in Kenya, 2SCALE farmers were between 37 and 42 percentage points more likely to have a crop under contract (Table 13). The programme did not seem to have an impact on measures of the diversity of crop production because the estimates for the number of crops produced, the Shannon index, and the Simpson index were not statistically different from zero. However, the lack of findings on crop diversity could correspond to farmers concentrating their production in Sorghum as a result of the 2SCALE partnership.

Table 13. Impact Estimates All Crops (Kenya)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Producing targeted crop: sorghum	0.43***	0.42***	0.92	0.98	0.98	0.61	1,326
	(13.37)	(13.37)					
No. of crops produced	-0.05	-0.02	3.18	3.37	3.80	4.05	1,326
	(-0.30)	(-0.12)					
Shannon index	-0.02	-0.01	0.99	1.08	1.02	1.13	1,326
	(-0.40)	(-0.23)					
Simpson index	0.01	0.01	1.43	1.38	1.42	1.37	1,326
	(0.50)	(0.32)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. Only panel observations are included. T = treatment group; C = comparison group.*p < .10. **p < .05. ***p < .01.

Additional Income-Related Outcomes

Because the partnership in Kenya developed a financial package for the farmers and because the partnership relationship was that of a client, we include tables that focus on credit and household income sources. Table 14 shows statistically significant increases in borrowing and debt from loans contracted in the past 12 months in Kenya, which we interpret as farmers being able to have access to credit thanks to the relationship with the programme.

Table 14. Nonfarm Business and Credit (Kenya)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Borrowed on credit from someone outside the household	0.08**	0.08**	0.21	0.13	0.26	0.10	1512
	(2.07)	(2.08)					

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
Debt from loans contracted in the past 12 months	3,438***	3,473***	3,371	2,004	6,029	1,224	1497
	(2.83)	(2.86)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Food Security

Table 15 shows that in Kenya, 2SCALE farmers improved across multiple subjective dimensions of food security, but a statistically significant decrease occurred in the likelihood of eating meat or fish five or more times in last month. However, progress was made in food security: treatment farmers were between 14 percentage points more likely to report being better off than 12 months ago, approximately 3 percentage points more likely to eat more than one meal a day, and 34 percentage points more likely to eat three or more meals a day. These increases corresponded to statistically significant decreases in our estimate of the food insecurity scale, suggesting that farmers have improved their food security.

Table 15. Self-Assessed Poverty (Kenya)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
Food consumption: more than adequate	0.02	0.02	0.23	0.12	0.21	0.07	1512
	(0.68)	(0.66)					
Does not consider itself very poor	-0.01	-0.01	0.98	0.88	0.99	0.89	1512
	(-0.34)	(-0.37)					
Better off than 12 months ago	0.13***	0.14***	0.46	0.44	0.53	0.37	1512
	(2.82)	(2.81)					
Eats more than one meal a day	0.03*	0.03*	0.98	0.97	0.99	0.96	1512
	(1.95)	(1.94)					
Eats three or more meals a day	0.34***	0.34***	0.75	0.77	0.92	0.59	1512
	(8.49)	(8.41)					
Ate meat or fish five or more times in the last month	-0.12***	-0.12***	0.32	0.17	0.13	0.10	1512
	(-3.15)	(-3.10)					
Food insecurity scale	-0.86*	-0.82*	1.55	5.02	2.09	6.42	1498
	(-1.93)	(-1.84)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

When examining the specific food items that farm households consumed in the last 7 days, Table 16 shows mixed results, with statistically significant increases for some food groups (e.g., treatment farmers were 4 percentage points more likely to consume vegetables, among others) and statistically significant decreases for others (e.g., treatment farmers were 24 percentage points less likely to consume roots or tubers, among others). Overall, these mixed results contributed to the FANTA measure of household dietary diversity being not statistically different from zero.

Table 16. Fanta Variables and Dietary Diversity Score (Kenya)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable:	(1)	(2)	(4)	(5)	(6)	(7)	*N*
In the last 7 days, household consumed grains or cereals	-0.00	-0.00	1.00	1.00	0.99	0.99	1512
	(-0.45)	(-0.40)					
In the last 7 days, household consumed roots or tubers	-0.24***	-0.24***	0.86	0.61	0.86	0.86	1512
	(-6.09)	(-6.16)					
In the last 7 days, household consumed vegetables	0.04***	0.04***	0.99	1.00	0.99	0.95	1512
	(2.93)	(2.93)					
In the last 7 days, household consumed fruits	0.05	0.05	0.81	0.65	0.91	0.71	1512
	(1.14)	(1.17)					
In the last 7 days, household consumed red meat or poultry	0.12**	0.12**	0.59	0.48	0.74	0.52	1512
	(2.40)	(2.53)					
In the last 7 days, household consumed eggs	0.17***	0.17***	0.62	0.59	0.69	0.49	1512
	(3.34)	(3.46)					
In the last 7 days, household consumed fish or shellfish	0.12***	0.12***	0.03	0.09	0.14	0.08	1512
	(4.18)	(4.18)					
In the last 7 days, household consumed legumes	0.03	0.03	0.98	0.96	0.95	0.90	1512
	(1.16)	(1.16)					
In the last 7 days, household consumed milk or milk products	-0.10***	-0.10***	0.95	0.82	0.81	0.78	1512
	(-2.79)	(-2.79)					
In the last 7 days, household consumed oils or fats	-0.14***	-0.13***	0.95	0.97	0.82	0.97	1512
	(-5.24)	(-5.22)					
In the last 7 days, household consumed sweets, sugar, or honey	0.06**	0.06**	0.97	0.95	0.93	0.86	1512
	(2.39)	(2.43)					
In the last 7 days, household consumed condiments	0.09***	0.09***	0.98	0.99	0.99	0.90	1512
	(4.89)	(4.88)					
Household Dietary Diversity Score	0.14	0.15	9.75	9.13	9.84	9.08	1505
	(0.90)	(0.99)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Non-independent Product Commercialisation

In the household survey, we asked questions related to whether farmers sold their crop independently, through non-independent production commercialisation methods, or a combination of both. Table 17 reports the impact estimates for questions related to non-independent product commercialisation methods in Kenya. Perhaps due to the fact that 2SCALE activities involved the development of a comprehensive financial package, treatment farmers were more likely to report that clients facilitated access to credit (although this was only weakly significant). However, given the model of training smallholders in Kenya through farmer field schools, it is surprising that we find that treatment farmers are less likely to report that clients facilitated access to agricultural training. Despite the mixed results on aspects that clients provided access, treatment farmers were 14 percentage points more likely to report that they would sell to the clients in the future.

Table 17. Non-independent Product Commercialisation (Kenya)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Client facilitated access to agricultural. Training	-0.10***	-0.10***	0.29	0.01	0.19	0.02	1512
	(-3.14)	(-3.12)					
Client facilitated access to: credit	0.03*	0.03*	0.03	0.00	0.06	0.00	1512
	(1.91)	(1.87)					
Would sell to client in future	0.14***	0.14***	0.59	0.08	0.68	0.04	1511
	(3.49)	(3.53)					

Notes: All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Uganda

In Uganda, 2SCALE worked with an association of cotton farmers through the Nyakatonzi Growers Cooperative Union. The partnership helped the cooperative procure and install oil milling machinery. However, we understand that there was a significant delay in setting up the machinery such that the mill was only recently established. Additionally, the partnership activities also involved agricultural and financial training, increasing access to credit, and developing the market for oil products and secondary products, although we understand that the pilot which was to develop and test cotton and soy-oil based products could not be implemented. We examine the impact of the programme on cultural practices, which may have been influenced by the training; cotton production and crop diversity; access to credit, which may have been

influenced by the financial training or increased access to credit; food security measures; and networks and social capital since the 2SCALE model focused on the cotton cooperative.

Parcels and Cultural Practices (Intermediate Outcomes)

Table 18 reports the impact estimates for using a variety of cultural practices in Uganda. We found a statistically significant decrease in the likelihood of planting certified varieties. Despite this finding, treatment farmers have farms that were 15 percentage points less likely to be affected by weeds and 12 percentage points less likely to report that their last harvest was affected by insects, a fungus, or disease, which can happen as a result of applying more appropriate cultural practices learned in the trainings.

Table 18. Parcels and Cultural Practices (Uganda)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Resistant varieties planted	-0.03	-0.04	0.46	0.38	0.50	0.45	1493
	(-0.51)	(-0.74)					
Certified varieties planted	-0.17***	-0.15***	0.78	0.62	0.77	0.78	1493
	(-3.62)	(-2.93)					
Residue removed prior to planting	-0.02	-0.02	0.87	0.76	0.86	0.76	1493
	(-0.38)	(-0.38)					
Crops planted early	-0.04	-0.02	0.87	0.77	0.83	0.78	1493
	(-1.04)	(-0.52)					
Farm suffered from weeds	-0.09**	-0.15***	0.76	0.68	0.80	0.81	1493
	(-2.14)	(-2.97)					
Last harvest affected by insects, a fungus, or disease	-0.05	-0.12**	0.75	0.61	0.80	0.71	1493
	(-1.08)	(-2.34)					
Percentage quantity affected by pests or disease	-0.01	-0.02	0.17	0.10	0.23	0.18	1496
	(-0.80)	(-1.12)					
Percentage quality or price affected by pests or disease	-0.01	-0.02	0.17	0.10	0.19	0.14	1496
	(-0.83)	(-1.37)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Target Crop Outcomes

In Uganda, little evidence exists for changes by 2SCALE farmers in their production of cotton (Table 19). There was no statistically significant difference between treatment and control farmers in the quantity or value of harvest or gross margins. The delay in establishing the oil mill

could have explained the lack of findings related to cotton production. We also understand that the treatment area experienced major civil unrest in 2016, which also could explain the lack of findings on production measures.

Table 19. Impact Estimates Target Crop (Uganda)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	N
Quantity of the total harvest (kilograms)	-0.07	-0.02	5.83	5.78	5.84	5.86	1217
	(-0.80)	(-0.24)					
Value of the total harvest	-0.08	-0.01	12.74	12.69	13.13	13.16	1212
	(-0.90)	(-0.09)					
Gross margins	-0.12	-0.01	12.40	12.48	12.81	13.01	1147
	(-1.01)	(-0.09)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. Only panel observations are included. T = treatment group; C = comparison group.*p < .10. **p < .05. ***p < .01.

All Crop Outcomes

In Uganda, we do not find statistically significant changes in the likelihood of producing the targeted crop or in measures of diversity of crop production (Table 20).

Table 20. Impact Estimates All Crops (Uganda)

	DiD estimates		Baseline means		Baseline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Producing targeted crop: cotton	0.05	0.06	0.86	0.97	0.74	0.79	1413
	(1.44)	(1.38)					
No. of crops produced	0.25*	0.28	2.85	2.91	3.44	3.25	1413
	(1.71)	(1.58)					
Shannon index	0.07*	0.08	0.84	0.92	0.88	0.89	1413
	(1.65)	(1.62)					
Simpson index	-0.03	-0.04	1.49	1.44	1.49	1.47	1413
	(-1.26)	(-1.36)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. Only panel observations are included. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Additional Income-Related Outcomes

Even though the programme's activities included financial training and access to credit, there does not seem to be an impact across nonfarm business and credit (Table 21), aside from a statistically significant decrease in the interest in borrowing.

Table 21. Nonfarm Business and Credit (Uganda)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Borrowed on credit from someone outside the household	0.01	-0.01	0.29	0.10	0.34	0.14	1496
	(0.13)	(-0.13)					
Debt from loans contracted in the past 12 months	2,996	38,093	218,563	15,225	241,153	34,818	1489
	(0.05)	(0.55)					
Would apply for a loan if certain he will get it	-0.13***	-0.13**	0.54	0.40	0.47	0.46	1496
	(-2.58)	(-2.26)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Food Security

Table 22 shows that in Uganda, little change occurred in the subjective dimensions of food security. Most measures were not statistically distinguishable from zero. However, a reduction of 8 percentage points occurred in the likelihood of 2SCALE farmers eating more than one meal a day, and weak evidence (at the 10% significance level) showed a statistically significant decrease in the likelihood of eating meat or fish five or more times in last month. There was no statistically significant change in the likelihood of eating three or more meals a day in the DiD model with covariates. Overall, no statistically significant change occurred in the food insecurity scale.

Table 22. Self-Assessed Poverty (Uganda)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Food consumption: more than adequate	-0.01	0.02	0.03	0.02	0.05	0.05	1496
	(-0.64)	(0.72)					
Does not consider itself very poor	-0.07	-0.03	0.74	0.51	0.78	0.62	1496
	(-1.48)	(-0.63)					

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
Better off than 12 months ago	-0.03	-0.01	0.32	0.16	0.37	0.24	1496
	(-0.63)	(-0.20)					
Eats more than one meal a day	-0.07***	-0.08***	0.97	0.96	0.88	0.94	1496
	(-2.74)	(-2.67)					
Eats three or more meals a day	-0.05	0.00	0.23	0.14	0.21	0.18	1496
	(-1.33)	(0.06)					
Ate meat or fish five or more times in the last month	-0.06*	-0.06*	0.08	0.08	0.08	0.13	1496
	(-1.90)	(-1.83)					
Food insecurity scale	0.33	0.20	7.83	9.88	7.38	9.11	1484
	(0.48)	(0.26)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

As in Kenya, when examining the specific food items that farm households consumed in the last 7 days in Uganda, Table 23 shows mixed results, with some food groups having statistically significant increases and some food groups having statistically significant decreases. Overall, these mixed results contributed to the FANTA measure of household dietary diversity being not statistically different from zero.

Table 23. Fanta Variables and Dietary Diversity Score (Uganda)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
In the last 7 days, household consumed cereals or grains, including millet and sorghum	-0.05**	-0.05**	0.97	0.95	0.93	0.96	1496
	(-2.23)	(-1.98)					
In the last 7 days, household consumed potatoes, yams, cassava, or other related foods	-0.05**	-0.04	0.92	0.92	0.94	0.99	1496
	(-2.21)	(-1.48)					
In the last 7 days, household consumed vegetables	-0.03	-0.03	0.90	0.85	0.95	0.93	1496
	(-0.89)	(-1.01)					
In the last 7 days, household consumed fruits	-0.10**	-0.08	0.79	0.52	0.72	0.54	1496
	(-2.06)	(-1.50)					
In the last 7 days, household consumed red meat or poultry	-0.01	0.04	0.53	0.39	0.62	0.48	1496
	(-0.12)	(0.68)					
In the last 7 days, household consumed eggs	-0.03	-0.01	0.25	0.16	0.30	0.24	1496
	(-0.59)	(-0.22)					
In the last 7 days, household consumed fresh or	-0.01	0.01	0.38	0.49	0.48	0.61	1496

Dependent variable	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
	(1)	(2)	(4)	(5)	(6)	(7)	*N*
dried fish or shellfish	(-0.25)	(0.17)					
In the last 7 days, household consumed beans, peas, lentils, or nuts	-0.01	-0.03	0.99	0.96	0.97	0.96	1496
	(-0.63)	(-1.34)					
In the last 7 days, household consumed milk, cheese, yogurt, or other milk products	-0.01	0.05	0.32	0.09	0.39	0.17	1496
	(-0.23)	(1.04)					
In the last 7 days, household consumed oils and fats	-0.11**	-0.14**	0.58	0.47	0.60	0.59	1496
	(-2.12)	(-2.39)					
In the last 7 days, household consumed sweets, sugar, or honey	-0.03	-0.02	0.82	0.71	0.71	0.64	1496
	(-0.73)	(-0.47)					
In the last 7 days, household consumed any other foods, such as condiments or coffee	0.07*	0.02	0.92	0.94	0.53	0.48	1496
	(1.79)	(0.44)					
Household Dietary Diversity Score	-0.46**	-0.37	8.37	7.47	8.15	7.70	1486
	(-1.97)	(-1.43)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. *** p<0.01

Networks and Social Capital

Table 24 reports some estimates of questions related to farmer networks, relationships, and social capital in Uganda, where 2SCALE operated through an association of cotton farmers. Treatment farmers were more likely to report that their group facilitated access to credit and processing, which is consistent with the activities that the partnership provided. Most likely due to the delay in establishing the oil mill, treatment farmers were less likely to report that their group facilitated access to machinery.

Table 24. Networks and Social Capital (Uganda)

Dependent variable	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Group facilitated access to: machinery	-0.06**	-0.06**	0.18	0.00	0.13	0.00	1496
	(-2.17)	(-2.00)					
Group facilitated access to: credit	0.03	0.07**	0.14	0.00	0.18	0.00	1496
	(1.11)	(2.14)					
Group facilitated access to: agr. training	-0.01	0.04	0.37	0.01	0.36	0.00	1496
	(-0.18)	(0.89)					
Group facilitated access to: marketing	0.03	0.01	0.18	0.00	0.21	0.00	1496
	(0.95)	(0.42)					

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Group facilitated access to: processing	0.03***	0.03**	0.01	0.00	0.04	0.00	1496
	(2.91)	(2.44)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Benin

In Benin, the lead partner EWIT provides trainings to vegetable farmers, increased access to finance, and an improved supply of high-quality seeds. The activities also attempted to reduce barriers on access to agricultural inputs, and trained women on entrepreneurship. As such, we examine the impact of the program on cultural practices, which may have been influenced by the training; vegetable production and crop diversity; access to credit, income from the targeted crop; food security measures; non-independent product commercialization; and networks and social capital.

Parcels and Cultural Practices (Intermediate Outcomes)

Table 25 shows that in general in Benin cultural practices were improved, and the fraction of the farms affected by pests or disease declined. Specifically, treatments farmers were 25 percentage points more likely to plant resistant varieties and 24 percentage points more likely to remove residue prior to planting. Treatment farmers were 18 percentage points less likely to have a farm affected by weeds and 19 percentage points less likely to have the last harvest affected by insects, a fungus, or disease. Treatment farmers also had statistically significant decreases in the fraction of the crop quantity (6 percentage points) and quality or price (4 percentage points) affected by pests or disease. The improvement in cultural practices may be due to the training that farmers received through the learning plots, field visits, and training sessions.

Table 25. Parcels and Cultural Practices (Benin)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Resistant varieties planted	0.25***	0.25***	0.46	0.42	0.58	0.29	1515
	(5.21)	(5.23)					
Certified varieties planted	0.02	0.02	0.24	0.06	0.27	0.07	1515
	(0.44)	(0.48)					
Residue removed prior to planting	0.24***	0.24***	0.78	0.78	0.72	0.48	1515
	(4.88)	(5.34)					

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
Crops planted early	0.02	0.02	0.72	0.60	0.71	0.57	1515
	(0.36)	(0.36)					
Farm suffered from weeds	-0.18***	-0.18***	0.59	0.45	0.62	0.67	1515
	(-3.25)	(-3.30)					
Last harvest affected by insects, a fungus, or disease	-0.19***	-0.19***	0.26	0.11	0.57	0.61	1515
	(-4.11)	(-4.12)					
Percentage quantity affected by pests or disease	-0.06***	-0.06***	0.02	0.01	0.07	0.12	1520
	(-6.68)	(-6.73)					
Percentage quality or price affected by pests or disease	-0.04***	-0.04***	0.02	0.01	0.07	0.09	1520
	(-5.03)	(-5.08)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Target Crop Outcomes

In Benin, we observed a decrease in the use of organic and inorganic fertiliser, the use of pesticide, the cost of seed planted, and the cost of paid plus family labour (Table 26). Despite the statistically significant reduction in costs, the DiD model with covariates does not suggest statistically significant changes in the quantity or value of production or gross margins. However, this decreased use of inputs and lack of change in production values could be due to the collapse of the Nigerian market in 2016, stemming from the devaluation of the Nigerian currency. Additionally, around that time there came increased measures to protect domestic industries. We understand that most of the treatment farmers sold to the Nigerian market; whereas the control group did not and that the Nigerian market collapse adversely affected farmers who sold on that market. Farmers affected in 2016 may have had difficulty reinvesting funds in farming activities in 2017.

Table 26. Impact Estimates Target Crop (Benin)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
Organic fertiliser used	-0.16***	-0.15***	0.77	0.03	0.85	0.27	1172
	(-3.33)	(-3.07)					
Inorganic fertiliser used	-0.39***	-0.38***	0.85	0.79	0.60	0.94	1172
	(-8.50)	(-8.42)					

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
Pesticide use	-0.16***	-0.17***	0.86	0.75	0.76	0.81	1172
	(-3.62)	(-3.63)					
Cost of seed planted	-1.55***	-1.55***	10.65	10.86	9.03	10.79	1151
	(-6.73)	(-6.22)					
Cost of pesticide	-0.18	-0.09	9.95	9.31	9.85	9.39	933
	(-1.36)	(-0.70)					
No. of family labour days	-2.78	-1.77	28.88	20.33	74.39	68.63	1146
	(-0.30)	(-0.19)					
No. of paid labour days	-4.34	2.52	35.86	3.71	44.80	16.99	1154
	(-0.60)	(0.31)					
Cost of paid labour	-0.06	0.00	10.52	10.05	10.82	10.41	723
	(-0.35)	(0.02)					
Cost of paid + family labour	-0.48***	-0.45***	11.36	11.11	11.96	12.20	1099
	(-3.63)	(-3.31)					
Quantity of the total harvest (kilograms)	-0.01	0.03	7.00	6.17	7.02	6.21	1127
	(-0.07)	(0.15)					
Value of the total harvest	-0.18	-0.16	12.98	12.04	12.63	11.87	1114
	(-0.98)	(-0.80)					
Gross margins	-0.46*	-0.41	12.89	11.84	12.78	12.18	683
	(-1.89)	(-1.61)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. Only panel observations are included. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

All Crop Outcomes

In Benin, we found statistically significant evidence of an increase in crop diversity (Table 27). Treatment farmers increased their number of crops produced by 0.71 crops, and the Shannon index was positive and statistically significant, suggesting an increase in the evenness and proportional abundance of crops. However, because the reverse was the case with the Simpson index (we found a decrease in that index, which suggests a decrease in the relative abundance of each crop), we should consider the findings on crop diversity with caution. Treatment farmers were also more likely to produce the targeted crop.

Table 27. Impact Estimates All Crops (Benin)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Producing targeted crop: tomatoes or chilies	0.21***	0.21***	0.97	0.97	0.76	0.56	1359
	(5.58)	(5.51)					
No. of crops produced	0.71***	0.71***	1.81	1.88	2.24	1.61	1359
	(7.89)	(7.62)					
Shannon index	0.37***	0.37***	0.52	0.59	0.61	0.32	1359
	(9.51)	(9.19)					
Simpson index	-0.23***	-0.23***	1.63	1.58	1.60	1.78	1359
	(-9.46)	(-9.20)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. Only panel observations are included. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Additional Income-Related Outcomes

In Benin, credit access seems to have improved: 2SCALE farmers experienced a statistically significant decrease (8 percentage points) in the probability of being turned down for borrowing outside the household (Table 28).

Table 28. Nonfarm Business and Credit (Benin)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Borrowed on credit from someone outside the household	-0.05	-0.05	0.32	0.03	0.32	0.09	1520
	(-1.65)	(-1.54)					
Debt from loans contracted in the past 12 months	-21,499	-21,922	178,252	7,640	196,411	47,299	1507
	(-0.49)	(-0.45)					
Tried to borrow from someone outside the household and turned down	-0.08***	-0.08***	0.11	0.00	0.14	0.10	1520
	(-2.87)	(-2.87)					
Would apply for a loan if certain he will get it	-0.02	-0.02	0.77	0.72	0.87	0.84	1520
	(-0.60)	(-0.56)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

In Benin, 2SCALE farmers had statistically significant decreases in the past year's contract farming income (Table 29), which may again be due to the Nigerian market collapse.

Table 29. Household Income Sources (Benin)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Past year's noncontract farming income	-15,121	-16,493	236,258	150,852	612,509	542,223	1504
	(-0.17)	(-0.19)					
Past year's contract farming income	-50,442**	-50,182**	87,591	10,035	30,690	3,577	1506
	(-2.48)	(-2.52)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Food Security

Table 30 shows that in Benin, a 9-percentage point increase occurred in the proportion of treatment households that reported their food consumption was more than adequate. There also was a 13-percentage point increase in the proportion of households that do not consider themselves to be very poor. On the other hand, we found a 13-percentage point decrease in the likelihood of eating meat or fish more than five times in the last month. We found an increase in the food insecurity scale, suggesting that households became more food insecure. However, these results on the food insecurity scale should be interpreted with caution because the specifications in columns 2 and 3 found statistically significant results that were opposite in sign.

Table 30. Self-Assessed Poverty (Benin)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Food consumption: more than adequate	0.09***	0.09***	0.10	0.05	0.26	0.12	1520
	(2.66)	(2.66)					
Does not consider itself very poor	-0.01	-0.01	0.88	0.79	0.97	0.89	1520
	(-0.36)	(-0.33)					
Better off than 12 months ago	-0.07	-0.07	0.71	0.41	0.52	0.28	1520
	(-1.38)	(-1.36)					
Eats more than one meal a day	0.02	0.02	0.98	0.95	0.98	0.93	1520
	(0.99)	(0.99)					
Eats three or more meals a day	0.05	0.05	0.82	0.73	0.76	0.61	1520
	(1.11)	(1.12)					
Ate meat or fish five or more times in last month	-0.13***	-0.13***	0.44	0.13	0.44	0.27	1520
	(-2.87)	(-2.84)					
Food insecurity scale	3.14***	3.13***	4.58	7.59	6.44	6.31	1520

Dependent variable	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
	(1)	(2)	(4)	(5)	(6)	(7)	*N*
	(4.50)	(4.46)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

When examining the specific food items that farm households consumed in the last 7 days, Table 31 shows mixed results, with statistically significant increases for some food groups and statistically significant decreases for others. Overall, these mixed results contributed to the FANTA measure of household dietary diversity being not statistically different from zero.

Table 31. Fanta Variables and Dietary Diversity Score (Benin)

Dependent variable	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
	(1)	(2)	(4)	(5)	(6)	(7)	*N*
In the last 7 days, household consumed grains or cereals	0.14***	0.14***	1.00	1.00	0.97	0.83	1520
	(6.58)	(6.61)					
In the last 7 days, household consumed roots or tubers	0.06**	0.06*	0.81	0.74	0.92	0.79	1520
	(1.97)	(1.96)					
In the last 7 days, household consumed vegetables	0.07***	0.07***	0.97	0.98	0.98	0.93	1520
	(3.70)	(3.72)					
In the last 7 days, household consumed fruits	-0.17***	-0.17***	0.69	0.48	0.92	0.89	1520
	(-5.02)	(-5.32)					
In the last 7 days, household consumed red meat or poultry	-0.01	-0.01	0.56	0.35	0.85	0.66	1520
	(-0.25)	(-0.22)					
In the last 7 days, household consumed eggs	-0.13***	-0.13***	0.53	0.20	0.72	0.53	1520
	(-2.93)	(-2.97)					
In the last 7 days, household consumed fish or shellfish	0.00	0.00	0.88	0.73	0.87	0.72	1520
	(0.05)	(0.05)					
In the last 7 days household consumed legumes	-0.05	-0.05	0.83	0.79	0.83	0.84	1520
	(-1.39)	(-1.39)					
In the last 7 days, household consumed milk or milk products	-0.03	-0.03	0.40	0.15	0.65	0.43	1520
	(-0.63)	(-0.65)					
In the last 7 days household consumed oils or fats	0.06*	0.06*	0.43	0.38	0.97	0.86	1520
	(1.65)	(1.71)					
In the last 7 days, household consumed sweets, sugar, or honey	0.20***	0.20***	0.93	0.88	0.83	0.59	1520
	(4.89)	(4.94)					

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
In the last 7 days, household consumed condiments	0.14***	0.14***	1.00	1.00	0.63	0.49	1520
	(3.88)	(3.89)					
Household Dietary Diversity Score	0.18	0.18	9.02	7.68	10.16	8.65	1515
	(0.83)	(0.85)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Non-independent Product Commercialisation

Table 32 reports the impact estimates for questions related to non-independent product commercialisation methods in Benin and suggests that treatment farmers had some challenges with client relationships. Surprisingly, treatment farmers were more likely to sell to a smaller number of clients. Treatment farmers sold to 0.25 fewer additional clients, which might be explained by treatment farmers' loss of their Nigerian clients due to the Nigerian market collapse in 2016. Treatment farmers were less likely to report that clients facilitated access to agricultural training or credit and were less likely to report that they would sell to the clients in the future.

Table 32. Non-independent Product Commercialisation (Benin)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
Number of clients sold to	-0.24***	-0.25***	0.32	0.00	0.11	0.03	1505
	(-5.65)	(-5.72)					
Clients facilitated access to agricultural training	-0.01*	-0.01*	0.01	0.00	0.00	0.00	1520
	(-1.95)	(-1.83)					
Clients facilitated access to credit	-0.02**	-0.02**	0.03	0.00	0.00	0.00	1520
	(-2.36)	(-2.35)					
Clients facilitated access to fertiliser	0.00	0.00	0.00	0.00	0.01	0.00	1520
	(0.61)	(0.61)					
Clients facilitated access to seeds	0.00	0.00	0.00	0.00	0.00	0.00	1520
	(0.00)	(0.01)					
Would sell to clients in future	-0.23***	-0.23***	0.29	0.00	0.09	0.02	1520
	(-9.64)	(-9.36)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Networks and Social Capital

Table 33 reports some estimates of questions related to farmer networks, relationships, and social capital in Benin. Treatment farmers in Benin reported some challenges regarding the aspects to which the producer groups facilitated access. Treatment farmers reported their producer groups were less likely to facilitate access to fertiliser, machinery, professional sprayers, labour, credit, and agricultural training.

Table 33. Networks and Social Capital (Benin)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Group facilitated access to: seeds	-0.00	-0.00	0.28	0.00	0.31	0.03	1520
	(-0.00)	(-0.00)					
Group facilitated access to: fertiliser	-0.11***	-0.11***	0.36	0.00	0.27	0.02	1520
	(-3.44)	(-3.26)					
Group facilitated access to: machinery	-0.06***	-0.06***	0.07	0.00	0.02	0.01	1520
	(-4.96)	(-4.94)					
Group facilitated access to: professional sprayers	-0.10***	-0.10***	0.11	0.00	0.01	0.00	1520
	(-8.41)	(-8.13)					
Group facilitated access to: labour	-0.05***	-0.05***	0.05	0.00	0.02	0.01	1520
	(-4.78)	(-4.69)					
Group facilitated access to: credit	-0.08***	-0.08***	0.14	0.00	0.06	0.00	1520
	(-3.38)	(-3.50)					
Group facilitated access to: agricultural training	-0.12***	-0.12***	0.23	0.00	0.11	0.00	1520
	(-6.12)	(-6.22)					
Group facilitated access to: marketing	0.01	0.01	0.02	0.00	0.02	0.00	1520
	(0.55)	(0.52)					
Group facilitated access to: distribution	-0.01	-0.01	0.02	0.00	0.01	0.00	1520
	(-1.45)	(-1.43)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*$p < .10$. **$p < .05$. ***$p < .01$.

Mali

In Mali, 2SCALE facilitated a partnership with SONAF, a trader of yellow maize that introduced high-yielding maize varieties and provided technical assistance regarding production, quality, and post-harvest management to smallholder farmers and improved farmers' access to credit. Accordingly, we examine the impact of the programme on cultural practices, which may have been influenced by the technical assistance; maize production and crop diversity; income from

the targeted crop; food security measures; as well as non-independent product commercialization and networks and social capital.

Parcels and Cultural Practices (Intermediate Outcomes)

Table 34 reports the impact estimates for using a variety of cultural practices in Mali. We found statistically significant increases in the likelihood of planting certified varieties and the likelihood of removing residue prior to planting. Despite being more likely to partake in these cultural practices, treatment farmers have farms that were 25 percentage points more likely to be affected by weeds; 15 percentage points more likely to report that their last harvest was affected by insects, a fungus, or disease; and were more likely to have a higher percentage of the quality or price of the harvest affected by pests or disease.

Table 34. Parcels and Cultural Practices (Mali)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Resistant varieties planted	0.04	0.05	0.29	0.31	0.55	0.53	1584
	(0.69)	(1.10)					
Certified varieties planted	0.19**	0.20***	0.50	0.70	0.75	0.75	1584
	(2.44)	(3.56)					
Residue removed prior to planting	0.19***	0.19***	0.49	0.58	0.76	0.65	1584
	(2.81)	(3.35)					
Crops planted early	0.03	0.03	0.44	0.45	0.45	0.43	1584
	(0.38)	(0.52)					
Farm suffered from weeds	0.25***	0.25***	0.18	0.26	0.68	0.50	1584
	(3.53)	(4.44)					
Last harvest affected by insects, a fungus, or disease	0.15**	0.15***	0.09	0.13	0.26	0.15	1584
	(2.39)	(2.90)					
Percentage quantity affected by pests or disease	0.00	0.00	0.01	0.01	0.01	0.00	1592
	(1.17)	(1.25)					
Percentage quality or price affected by pests or disease	0.01**	0.01**	0.01	0.01	0.01	0.01	1592
	(2.31)	(2.49)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10 **p < .05. ***p < .01.

Target Crop Outcomes

In Mali, Table 35 shows that the changes in the quantity of the total harvest, the value of total harvest, and gross margins were not statistically distinguishable from zero. The lack of findings related to production of the target crop may be due to the nature of the beneficiary farmers included in the study. The farmers in the treatment group came from two of the larger producer organizations (PO) that we understand happened to have difficulty reimbursing the credit for inputs obtained through SONAF and also experienced governance challenges. 2SCALE

representatives reported that the smaller POs that they work with in Mali tended to be more efficient, and the results of this evaluation do not apply to the partnership model that uses the smaller POs.

Table 35. Impact Estimates Target Crop (Mali)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Quantity of the total harvest (kilograms)	-0.03	-0.02	5.78	5.87	6.51	6.63	828
	(-0.19)	(-0.13)					
Value of the total harvest	0.02	0.07	11.07	11.14	11.21	11.26	686
	(0.11)	(0.59)					
Gross margins	0.09	0.15	10.80	10.91	10.96	10.98	643
	(0.36)	(0.86)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. Only panel observations are included. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

All Crop Outcomes

Interestingly, in Mali, 2SCALE farmers were 6 percentage points less likely to have produced the targeted crop (Table 36). However, this finding could result from a higher share of farmers in the treatment group producing local or improved maize at baseline versus the comparison group. The already high penetration of treatment farmers producing maize provided less opportunity to increase this fraction in the treatment group. As with Benin, we found statistically significant evidence of an increase in crop diversity. Treatment farmers increased their number of crops produced by 0.25 crops, and the Shannon index was positive and statistically significant, suggesting an increase in the evenness and proportional abundance of crops. However, because the reverse was the case with the Simpson index (we found a decrease in that index, which suggests a decrease in the relative abundance of each crop), we should consider the findings on crop diversity with caution.

Table 36. Impact Estimates All Crops (Mali)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Producing targeted crop:	-0.08**	-0.06**	0.96	0.90	0.94	0.95	891
local and improved maize	(-2.24)	(-2.23)					
No. of crops produced	0.23***	0.25***	1.24	1.45	1.85	1.83	891
	(4.07)	(4.08)					
Shannon index	0.22***	0.21***	0.05	0.16	0.28	0.17	891
	(5.73)	(5.70)					

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Simpson index	-0.15***	-0.15***	1.96	1.89	1.80	1.88	891
	(-5.71)	(-5.70)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. Only panel observations are included. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Additional Income-Related Outcomes

Mali had a statistically significant decrease in the past year's noncontract farming income measure (Table 37), which may be due to the difficulty in reimbursing credit or PO governance issues.

Table 37. Household Income Sources (Mali)

	DiD estimates		Baseline means		Endline means		
Dependent variable	No Covs (1)	Covs (2)	T (4)	C (5)	T (6)	C (7)	*N*
Past year's noncontract farming income	-110,019***	-113,849***	72,050	78,736	206,542	323,248	1581
	(-3.12)	(-2.90)					
Past year's contract farming income	-23,942	-26,291	5,860	5,343	91,020	114,445	1577
	(-0.82)	(-0.89)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Food Security

Table 38 shows that in Mali, 2SCALE farmers improved across some subjective dimensions of food security (e.g., the proportion of households who consider their food consumption to be more than adequate and the proportion of households who eat more than one meal a day), but a statistically significant decrease occurred in the likelihood of eating three or more meals a day, which may have contributed to the weak evidence (at the 10% significance level) of an increase in the overall food insecurity scale.

Table 38. Self-Assessed Poverty (Mali)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
Food consumption: more than adequate	0.11***	0.12***	0.22	0.29	0.25	0.20	1598
	(2.68)	(2.72)					
Household not very poor	0.01	0.02	0.96	0.97	1.00	1.00	1598
	(1.06)	(1.27)					
Better off than 12 months ago	0.05	0.05	0.72	0.76	0.71	0.70	1598
	(1.13)	(1.15)					
Eats more than one meal a day	0.05**	0.06**	0.82	0.88	0.97	0.99	1598
	(2.07)	(2.09)					
Eats three or more meals a day	-0.07*	-0.07*	0.64	0.69	0.85	0.97	1598
	(-1.77)	(-1.73)					
Ate meat or fish five or more times in the last month	0.03	0.02	0.02	0.03	0.45	0.43	1598
	(0.69)	(0.65)					
Food insecurity scale	1.05**	0.98*	1.82	2.00	5.84	4.98	1585
	(2.03)	(1.89)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

When examining the specific food items that farm households consumed in the last 7 days, Table 39 shows that there was little change across food groups or the household dietary diversity measure with the DiD model with covariates.

Table 39. Fanta Variables and Dietary Diversity Score (Mali)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	**(1)**	**(2)**	**(4)**	**(5)**	**(6)**	**(7)**	***N***
In the last 7 days, household consumed grains and cereals	-0.08	-0.08*	0.99	0.92	0.97	0.98	1592
	(-1.40)	(-1.66)					
In the last 7 days, household consumed roots and tubers	-0.00	-0.00	0.80	0.75	0.92	0.86	1592
	(-0.13)	(-0.06)					
In the last 7 days. household consumed vegetables	0.09	0.09	0.90	0.84	0.85	0.70	1592
	(1.31)	(1.44)					

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
In the last 7 days, household consumed fruits	-0.00	-0.01	0.71	0.60	0.78	0.67	1592
	(-0.05)	(-0.19)					
In the last 7 days, household consumed red meat or poultry	0.03	0.03	0.74	0.74	0.98	0.95	1592
	(0.30)	(0.42)					
In the last 7 days, household consumed eggs	0.02	0.02	0.61	0.54	0.63	0.54	1592
	(0.37)	(0.37)					
In the last 7 days, household consumed fish or shellfish	0.12	0.12*	0.65	0.76	0.81	0.80	1592
	(1.42)	(1.82)					
In the last 7 days, household consumed legumes	-0.04	-0.03	0.90	0.81	0.80	0.75	1592
	(-0.63)	(-0.64)					
In the last 7 days household consumed milk or milk products	0.15**	0.15***	0.74	0.79	0.85	0.76	1592
	(2.52)	(2.88)					
In the last 7 days household consumed oils or fats	-0.05	-0.06	0.38	0.29	0.72	0.69	1592
	(-1.15)	(-1.19)					
In the last 7 days, household consumed sweets, sugar, or honey	-0.02	-0.02	0.86	0.83	0.97	0.96	1592
	(-0.67)	(-0.96)					
In the last 7 days, household consumed condiments	-0.10	-0.10	0.96	0.84	0.98	0.96	1592
	(-1.09)	(-1.43)					
Household Dietary Diversity Score	0.12	0.11	9.23	8.70	10.28	9.62	1592
	(0.21)	(0.25)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Non-independent Product Commercialisation

Table 40 reports the impact estimates for questions related to non-independent product commercialisation methods in Mali. There were no statistically significant changes in the number of clients sold to, the likelihood that clients facilitated access to agricultural training, or farmer's willingness to sell to clients again in the future.

Table 40. Non-independent Product Commercialisation (Mali)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Number of clients sold to	0.01	0.01	0.07	0.05	0.21	0.18	1588
	(0.30)	(0.32)					
Clients facilitated access to agricultural training	-0.00	-0.00	0.00	0.00	0.09	0.09	1592
	(-0.12)	(-0.15)					
Would sell to clients in future	0.01	0.01	0.03	0.01	0.17	0.14	1592
	(0.25)	(0.19)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Networks and Social Capital

Table 41 reports some estimates of questions related to farmer networks, relationships, and social capital in Mali. There were no statistically significant changes in treatment farmers' likelihood of reporting that their producer groups facilitated access to training, transportation, marketing, processing, or distribution.

Table 41. Networks and Social Capital (Mali)

	DiD estimates		Baseline means		Endline means		
	No Covs	Covs	T	C	T	C	
Dependent variable	(1)	(2)	(4)	(5)	(6)	(7)	*N*
Group facilitated access to: agricultural training	-0.04	-0.04	0.04	0.01	0.17	0.17	1592
	(-1.06)	(-1.06)					
Group facilitated access to: marketing	0.00	0.00	0.00	0.00	0.01	0.01	1592
	(0.37)	(0.38)					
Group facilitated access to: processing	-0.00	-0.00	0.00	0.00	0.00	0.00	1592
	(-1.00)	(-1.00)					
Group facilitated access to: distribution	-0.00	-0.01	0.00	0.00	0.00	0.00	1592
	(-1.33)	(-1.35)					

Note. All estimates use DiD modelling. The estimates in column 1 do not include control variables. The estimates in column 2 include control variables. T-statistics robust to heterogeneity are presented in parentheses. T = treatment group; C = comparison group.
*p < .10. **p < .05. ***p < .01.

Ghana

Recall that for Ghana, no pure baseline exists because the programme had been in operation across both treatment and comparison groups since 2013. Thus, we used a different identification strategy that combines regression and propensity score matching methods using only the endline data. We have included a detailed description of the methodology for Ghana in Appendix B. Accordingly, the format of the tables differs in that Column 1 includes the impact estimate, and we only report the endline means.

Target Crop Outcomes

In Ghana, evidence exists of some statistically significant changes in input use (a decrease in the use of pesticide, the cost of seed planted, the cost of pesticide, and the cost of paid and family labour and an increase in the use of inorganic fertiliser and the cost of paid labour) and in farmers' decreased production of soybeans (Table 42). 2SCALE farmers experienced statistically significant decreases in the quantity of the total harvest (24%), the value of the total harvest (25%), and the gross margin (31%).

Table 42. Impact Estimates Target Crop (Ghana)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)	N
Organic fertilizer used	0.00	0.01	0.01	763
	(0.01)			
Inorganic fertilizer used	0.04***	0.01	0.05	763
	(0.01)			
Pesticide use	-0.31***	0.67	0.36	763
	(0.04)			
Cost of seed planted	-0.42***	3.76	3.34	763
	(0.05)			
Cost pesticide	-0.12*	3.71	3.60	382
	(0.07)			
No. family labour days	-0.67	17.77	17.10	760
	(1.26)			
No. paid labour days	0.42	4.24	4.67	763
	(0.45)			
Cost paid labour	0.30***	4.30	4.60	523
	(0.08)			
Cost paid + family labour	-0.12**	5.85	5.72	756
	(0.06)			
Quantity of the total harvest (kilograms)	-0.24***	5.52	5.27	757
	(0.06)			
Value of total harvest	-0.25***	5.92	5.68	756
	(0.06)			
Gross margins	-0.31***	5.67	5.36	715
	(0.08)			

Note. Impacts are estimated using the inverse-probability-weighted regression adjustment. Robust t-statistics in parentheses. All estimations control for HH Size at baseline, dummies for household age composition; main respondent's gender, age, marital status, years of farming experience, education, use of computer, internet, and mobile phone; distance to centre of town; dummies for most important parcel being black/dark coloured, very fertile, flat or slightly flat, and has no erosion; number of rooms; dwelling is owner-occupied; shelter is traditional House; dummies for type of walls; dummies for type of floors; dummies for types of water access; household has electricity; dummies for types of toilet; and asset index. * p<0.10; ** p<0.05; *** p<0.01.

All Crop Outcomes

As with Mali and Benin, we found statistically significant evidence of an increase in crop diversity (Table 43). Treatment farmers increased their number of crops produced by 0.22 crops, and the Shannon index was positive and statistically significant, suggesting an increase in the evenness and proportional abundance of crops. However, because the reverse was the case with the Simpson index (we found a decrease in that index, which suggests a decrease in the relative abundance of each crop), we should consider the findings on crop diversity with caution.

Table 43. Impact Estimates All Crops (Ghana)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)	N
Producing targeted crop: soybeans	-0.01	1.00	0.99	748
	(0.01)			
No. crops produced	0.22***	2.00	2.22	748
	(0.05)			
Shannon index	0.08***	0.61	0.69	748
	(0.03)			
Simpson index	-0.04***	1.59	1.54	748
	(0.02)			

Note. Impacts are estimated using the inverse-probability-weighted regression adjustment. Robust t-statistics in parentheses. All estimations control for HH Size at baseline, dummies for household age composition; main respondent's gender, age, marital status, years of farming experience, education, use of computer, internet, and mobile phone; distance to centre of town; dummies for most important parcel being black/dark coloured, very fertile, flat or slightly flat, and has no erosion; number of rooms; dwelling is owner-occupied; shelter is traditional House; dummies for type of walls; dummies for type of floors; dummies for types of water access; household has electricity; dummies for types of toilet; and asset index. * p<0.10; ** p<0.05; *** p<0.01

Additional Income-Related Outcomes

In Ghana the programme contributed to statistically significant increases in borrowing and debt from loans contracted in the past 12 months (Table 44), but treatment farmers were less likely to say they would apply for a loan if they were certain they could get it.

Table 44. Nonfarm Business and Credit (Ghana)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)	N
Borrowed on credit from someone outside the household	0.06**	0.07	0.13	776
	(0.03)			

Debt from loans contracted in the past 12 months	0.11**	0.14	0.25	776
	(0.05)			
Tried to borrow from someone outside the household and turned down	0.01	0.01	0.02	776
	(0.01)			
Would apply for a loan if certain he will get it	-0.12***	0.41	0.29	776
	(0.04)			

Note. Impacts are estimated using the inverse-probability-weighted regression adjustment. Robust t-statistics in parentheses. All estimations control for HH Size at baseline, dummies for household age composition; main respondent's gender, age, marital status, years of farming experience, education, use of computer, internet, and mobile phone; distance to centre of town; dummies for most important parcel being black/dark coloured, very fertile, flat or slightly flat, and has no erosion; number of rooms; dwelling is owner-occupied; shelter is traditional House; dummies for type of walls; dummies for type of floors; dummies for types of water access; household has electricity; dummies for types of toilet; and asset index. * p<0.10; ** p<0.05; *** p<0.01

Food Security

Table 45 shows that in Ghana, 2SCALE farmers improved across a couple of subjective dimensions of food security; however, no change occurred in the food insecurity measure. Nevertheless, progress was made in food security because treatment farmers were approximately 4 percentage points more likely to eat more than one meal a day and 7 percentage points more likely to eat meat or fish more than five times in the past month.

Table 45. Self-Assessed Poverty (Ghana)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)	N
Food consumption: more than adequate	0.02	0.12	0.14	776
	(0.03)			
Does not consider itself very poor	-0.03	0.93	0.90	776
	(0.03)			
Better off than 12 months ago	0.05	0.26	0.31	776
	(0.04)			
Eats more than one meal a day	0.04*	0.91	0.94	776
	(0.02)			
Eats three or more meals a day	-0.05	0.79	0.74	776
	(0.03)			
Ate meat or fish five or more times in last month	0.07***	0.06	0.13	776
	(0.03)			
Food insecurity scale	-0.07	2.94	2.87	776
	(0.38)			

Note. Impacts are estimated using the inverse-probability-weighted regression adjustment. Robust t-statistics in parentheses. All estimations control for All estimations control for HH Size at baseline, dummies for household age composition; main respondent's gender, age, marital status, years of farming experience, education, use of computer, internet, and mobile phone; distance to centre of town; dummies for most important parcel being black/dark coloured, very fertile, flat or slightly flat, and has no erosion; number of rooms; dwelling is owner-occupied; shelter is traditional House; dummies for type of walls; dummies for type of floors; dummies for types of water access; household has electricity; dummies for types of toilet; and asset index. *p < .10. **p < .05. ***p < .01.

Although Table 46 shows statistically significant changes for some food groups, overall there was no statistically significant impact on the FANTA measure of household dietary diversity.

Table 46. Fanta Variables and Dietary Diversity Score (Ghana)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)	N
In the last 7 days, household consumed cereals or grains, including millet and sorghum	-0.03**	0.98	0.95	776
	(0.01)			
In the last 7 days, household consumed potatoes, yams, cassava or other related foods	0.00	0.95	0.95	776
	(0.02)			
In the last 7 days, household consumed vegetables	0.05	0.84	0.89	776
	(0.03)			
In the last 7 days, household consumed fruits	0.18***	0.40	0.58	776
	(0.04)			
In the last 7 days, household consumed red meat or poultry	-0.08**	0.74	0.66	776
	(0.04)			
In the last 7 days, household consumed eggs	-0.02	0.40	0.38	776
	(0.04)			
In the last 7 days, household consumed fresh or dried fish or shellfish	-0.02	0.89	0.87	776
	(0.03)			
In the last 7 days household consumed beans, peas, lentils, or nuts	0.02	0.82	0.84	776
	(0.03)			
In the last 7 days, household consumed milk, cheese, yogurt, or other milk products	-0.10**	0.52	0.41	776
	(0.04)			
In the last 7 days, household consumed oils and fats	-0.05	0.79	0.74	776
	(0.04)			
In the last 7 days, household consumed sweets, sugar, or honey	-0.07**	0.92	0.86	776
	(0.03)			
In the last 7 days, household consumed any other foods, such as condiments or coffee	-0.02	0.66	0.64	776
	(0.04)			
Household Dietary Diversity Score	-0.15	8.91	8.76	776
	(0.20)			

Note. Impacts are estimated using the inverse-probability-weighted regression adjustment. Robust t-statistics in parentheses. All estimations control for All estimations control for HH Size at baseline, dummies for household age composition; main respondent's gender, age, marital status, years of farming experience, education, use of computer, internet, and mobile phone; distance to centre of town; dummies for most important parcel being black/dark coloured, very fertile, flat or slightly flat, and has no erosion; number of rooms; dwelling is owner-occupied; shelter is traditional House; dummies for type of walls; dummies for type of floors; dummies for types of water access; household has electricity; dummies for types of toilet; and asset index. *p < .10. **p < .05. ***p < .01

Networks and Social Capital

Table 47 reports some estimates of questions related to farmer networks, relationships, and social capital in Ghana. The treatment group's producer groups in Ghana seemed to facilitate access to a variety of production aspects. Specifically, 2SCALE increased the likelihood that treatment farmers' producer groups facilitated access to seeds, fertiliser, machinery, professional sprayers, credit, and marketing.

Table 47. Networks and Social Capital (Ghana)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)	N
Group facilitated access to: seeds	0.22***	0.07	0.29	776
	(0.03)			
Group facilitated access to: fertiliser	0.04**	0.02	0.05	776
	(0.02)			
Group facilitated access to: machinery	0.10***	0.06	0.16	776
	(0.03)			
Group facilitated access to: professional sprayers	0.03***	0.01	0.04	776
	(0.01)			
Group facilitated access to: labour	0.01	0.11	0.12	776
	(0.03)			
Group facilitated access to: credit	0.09***	0.04	0.13	776
	(0.02)			
Group facilitated access to: marketing	0.09***	0.03	0.12	776
	(0.02)			

Note. Impacts are estimated using the inverse-probability-weighted regression adjustment. Robust t-statistics in parentheses. All estimations control for All estimations control for HH Size at baseline, dummies for household age composition; main respondent's gender, age, marital status, years of farming experience, education, use of computer, internet, and mobile phone; distance to centre of town; dummies for most important parcel being black/dark coloured, very fertile, flat or slightly flat, and has no erosion; number of rooms; dwelling is owner-occupied; shelter is traditional House; dummies for type of walls; dummies for type of floors; dummies for types of water access; household has electricity; dummies for types of toilet; and asset index. *p < .10. **p < .05. ***p < .01.

Descriptive examination of awareness, use, and preferences

To examine awareness, use, and preference for the programme, we provide some graphical analyses of key factors related to programme participation. For this analysis, we restrict the sample to farmers only in the treatment group because the answers reported by the treatment group will be most informative for thinking about the relationship between farmers and 2SCALE. Our examination is descriptive in nature and cannot be used to infer causal relationships.

Of which components of the 2SCALE programme are farmers aware?

Data that describe to whom treatment farmers sold, the types of contracts the farmers had, the decision making between farmers and clients, and a breakdown of the activities to which clients or producer groups provided access provides insight into farmers' awareness of 2SCALE.

Figure 7 shows the breakdown of whether treatment farmers sold their crop all independently; sold all to an agent, broker, or company; or sold it part independently and part to an agent. Given that these graphs are restricted to treatment farmers only, we would have expected the fraction sold to an agent, broker, or company to be the dominant fraction, as is the case in Kenya, Uganda, and Ghana; but it is surprising that such a high fraction of farmers reported selling all the crop independently in Benin and Mali.

Figure 7. Method of Selling Target Crop

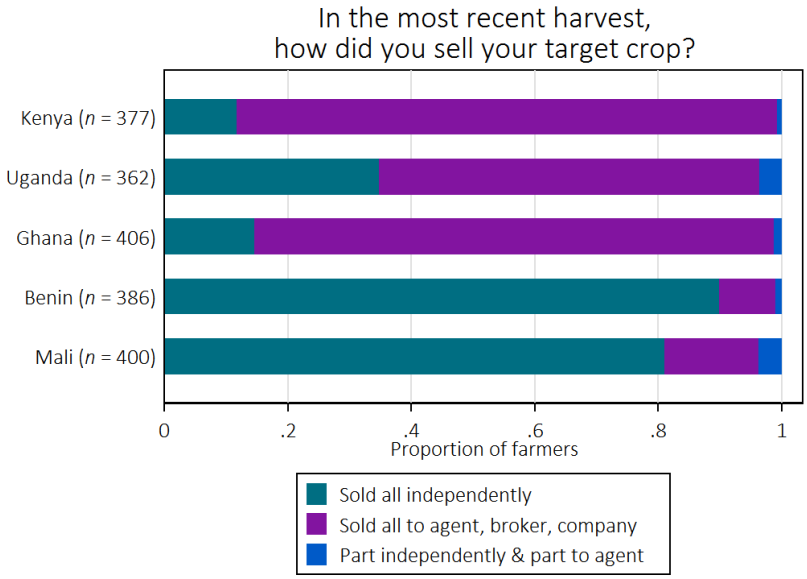
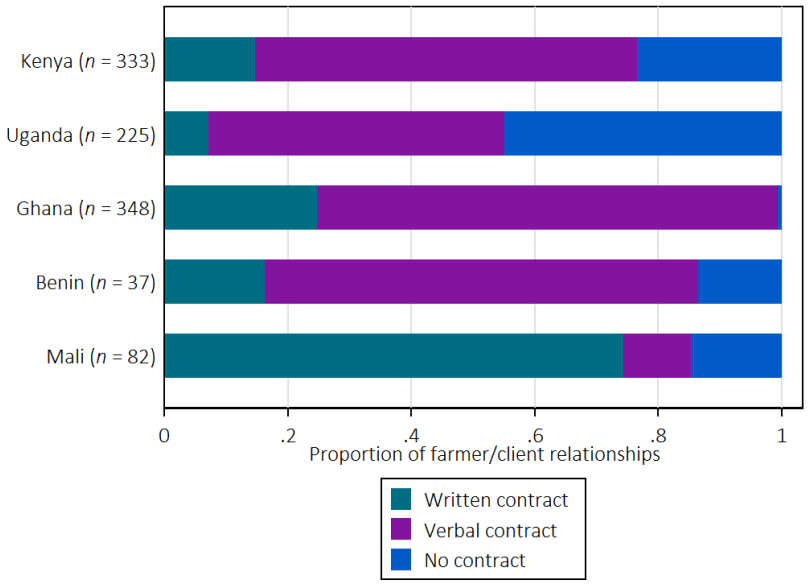


Figure 8 shows that, as we would expect, across the five countries, most farmers who sold non-independently had either a written or verbal contract; the exception was in Uganda, where more than 40% of the treatment farmers reported no contract.

Figure 8. Types of Contracts Used to Sell Target Crop



Looking at decision making, we rely on questions asked to farmers who sell through non-independent product commercialisation methods regarding decision making of key agricultural activities. For the various activities, farmers could report that they made the decision, the clients made the decision, or both parties made the decision.[6] We can think of a farmer's perceptions on who is involved in decision making as indicators for awareness because these data demonstrate whether farmers are aware of the relationships they have with clients.

Figures 9–13 show that across all five countries, among those farmers who sold non-independently, farmers were overwhelmingly responsible for farming decisions. In Kenya and Benin, clients tended to be involved in the decision making for crops, seeds, and the harvest delivery date. In Uganda and Ghana, the clients or both parties tended to be more involved in decision making across farming activities. In Mali, clients seemed to be infrequently involved in decision making.

[6] Responses of "No decision made" correspond to "Doesn't know."

Figure 9. Decision Making in Kenya

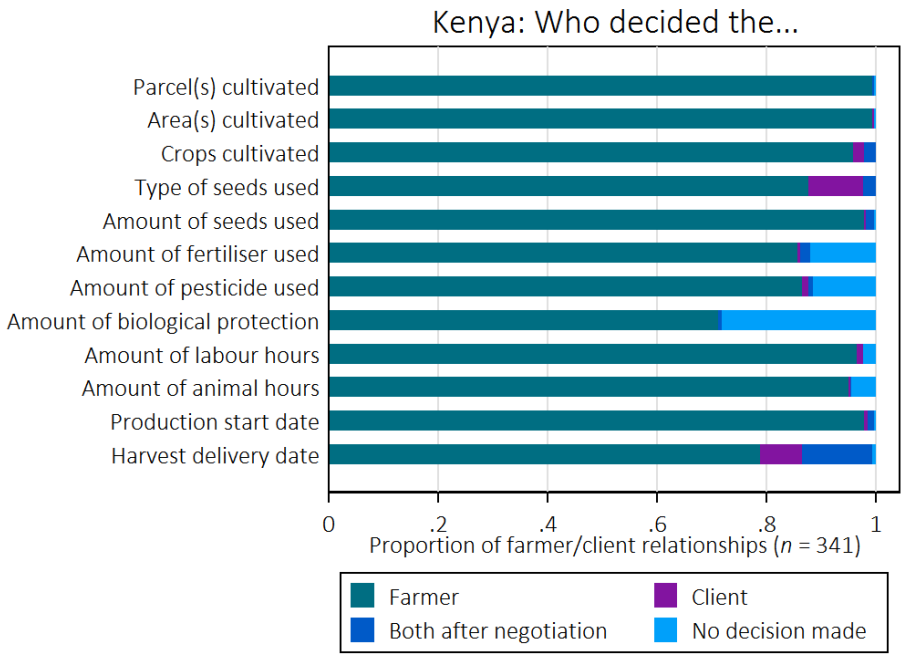


Figure 10. Decision Making in Uganda

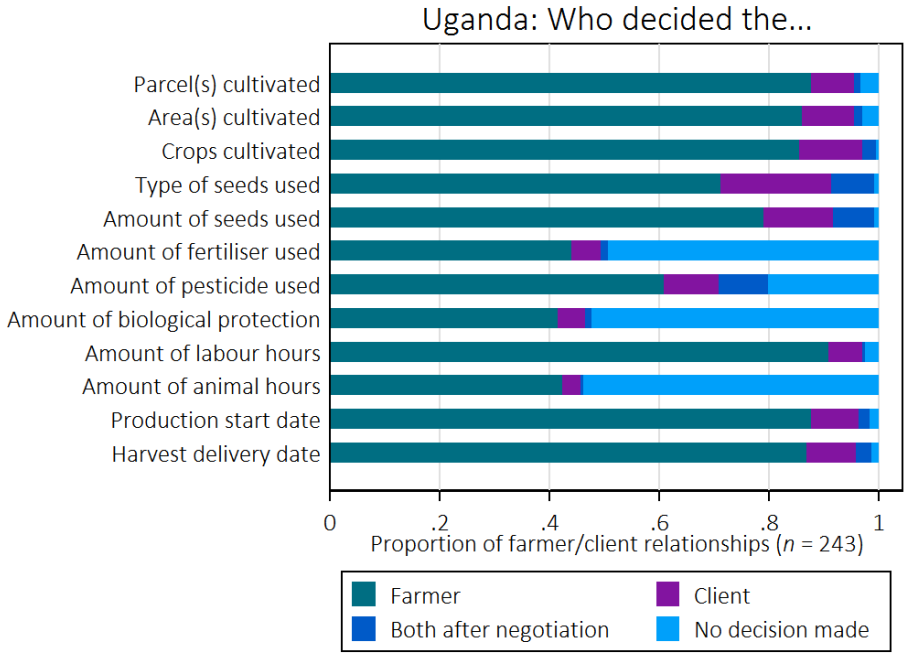


Figure 11. Decision Making in Ghana

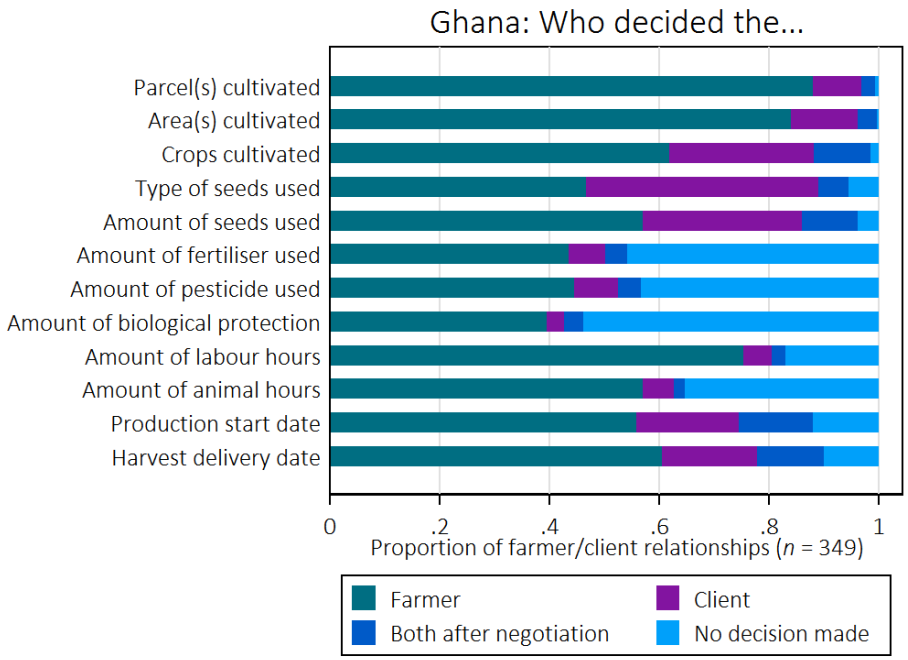


Figure 12. Decision Making in Benin

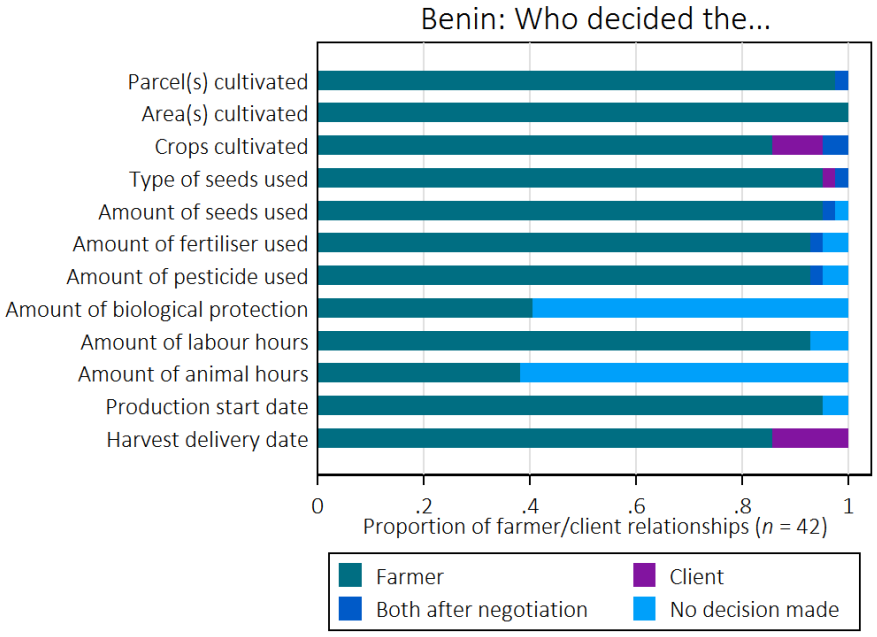
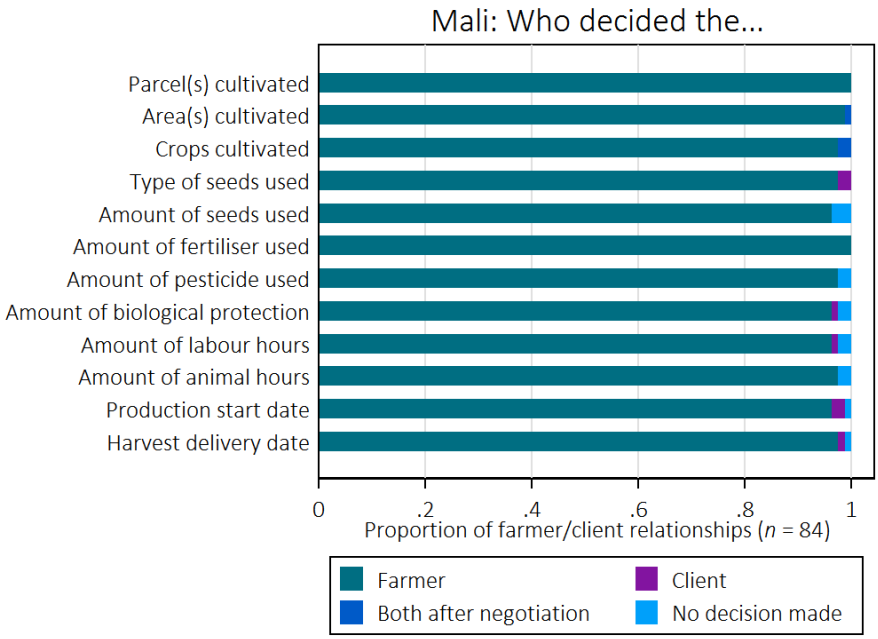


Figure 13. Decision Making in Mali



The inputs to which clients and producer groups facilitated access are displayed in Figures 14 and 15. For farmers selling non-independently, Figure 14 displays the percentages of farmer-client relationships that reported activities to which clients facilitated access, and Figure 15 displays the percentages based on farmers who reported belonging to a producer group. In most of the countries, treatment farmers reported a variety of inputs to which clients and producer groups provided access. Producer group provision of agricultural trainings was reported by a somewhat substantial fraction across all countries, and client provision of agricultural trainings followed a similar pattern, except in Benin where it was not reported. Provision of seed and credit also was common across all countries and across both client and producer group provision. Although some level of fertiliser provision was reported in all countries, this was more commonly reported in Mali.

Figure 14. Inputs to Which Client Facilitated Access

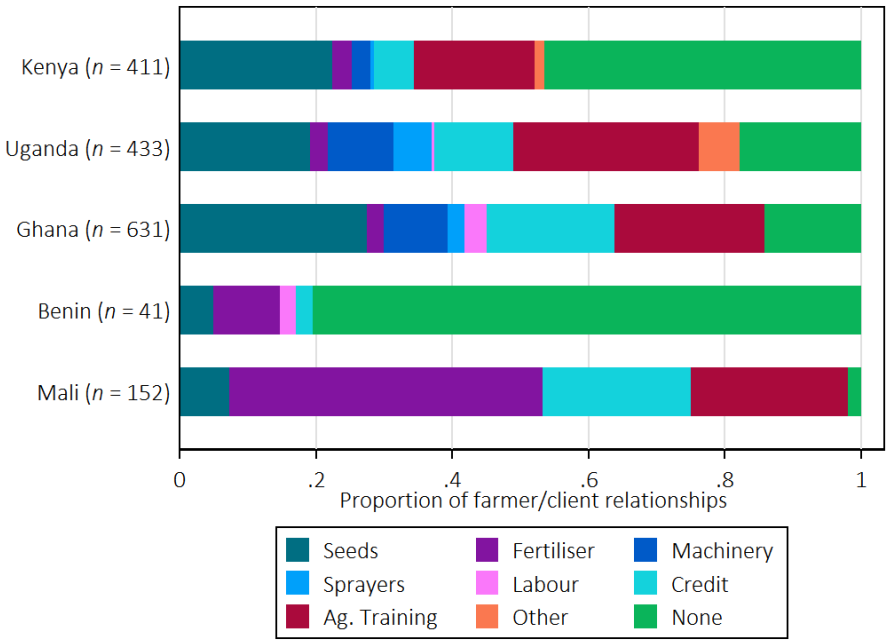
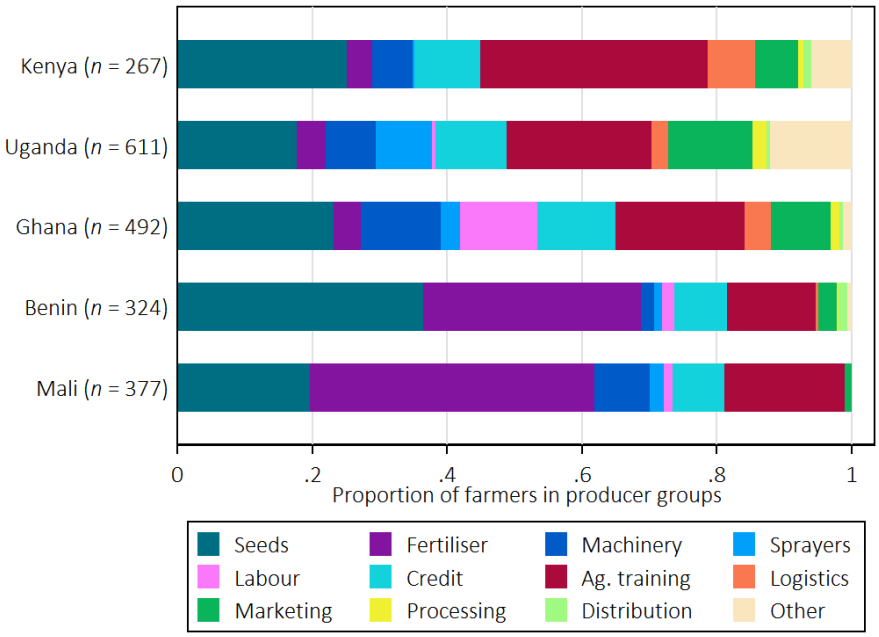


Figure 15. Inputs to Which Producer Group Facilitated Access



Which components of the 2SCALE programme have farmers used?

To examine what 2SCALE components treatment farmers used, we examine the activities that farmers used (corresponding to the activities that farmers indicated clients facilitated access to

from Figure 14), the number of times treatment farmers interacted with a representative from the producer group, the knowledge farmers had on key production values (price) during the production cycle, and the problems that 2SCALE farmers had when selling to clients.

Figure 16 shows that the components of 2SCALE that farmers used are relatively more concentrated than the activities to which the programme provides access (from Figure 14).[7] Agricultural trainings were commonly reported across all countries. All countries also reported using the access to credit and machinery, although these tended to be less commonly reported than agricultural trainings. More than 20% of the relationships used the provision of seeds in all countries except for Mali, which reported a smaller share of relationships using seed provision. Fertiliser was commonly used in Benin and Mali. Labour and sprayers were reported in all countries except for Mali.

Figure 16. Which Inputs/Activities Used

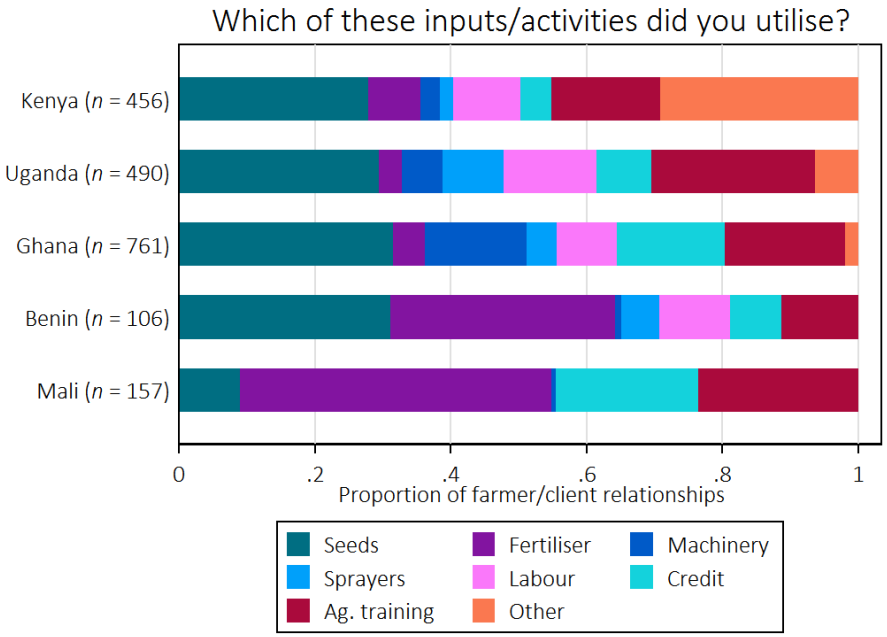


Figure 17 reveals that across countries, variation existed in the frequency with which treatment farmers, who belonged to producer groups, interacted with a representative from the producer group. However, in all the countries except Uganda, more frequent interaction was common. All countries reported weekly or monthly interactions for more than 45% of the those farmers.

[7] Farmers could select all that apply, which accounts for the numbers being greater than 400.

Figure 17. Farmers' Interaction With Producer Group Representatives

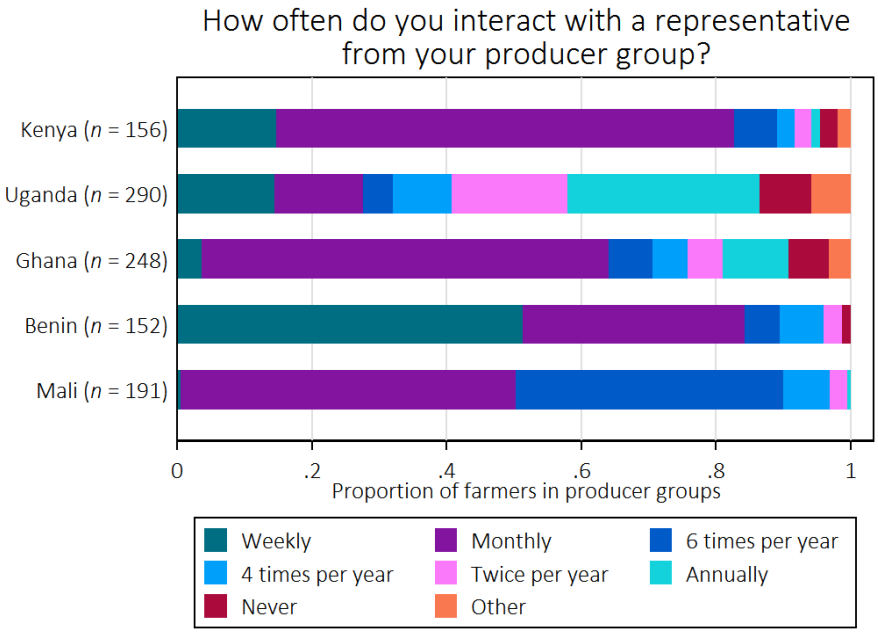
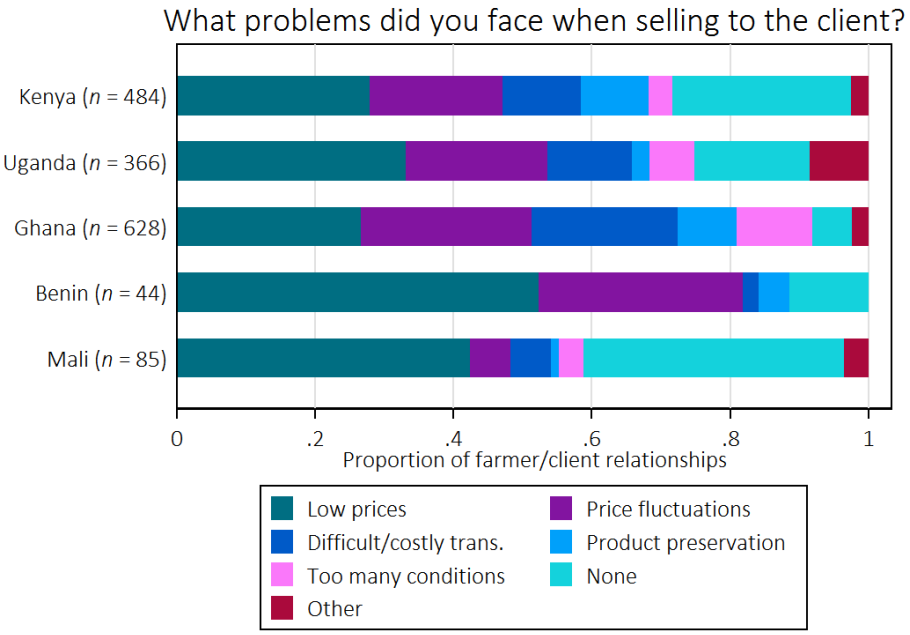


Figure 18 shows that low prices or price fluctuations were the most commonly reported problems in selling to clients across farmer/client relationships in all countries. Other problems that farmers reported included difficult or costly transaction costs, product preservation, or having too many conditions.

Figure 18. Non-independent Selling Problems



Which components of the 2SCALE programme do farmers prefer?

To examine what 2SCALE components treatment farmers prefer, we examine the main reasons why farmers did business with clients, the reasons that farmers belonged to producer groups, whether farmers would want to sell to clients again, and the reasons farmers gave for wanting to sell or not sell to clients again in the future.

Figures 19 and 20 show that farmers had diverse reasons for doing business with a client or belonging to a producer group. The variety in reported reasons for wanting to do business with a client or belong to a producer group lends support to 2SCALE's model of working with various stakeholders to identify the mechanisms that strengthen incentives for co-investment, connectedness, and competitiveness.

Figure 19. Main Reasons for Doing Business With the Client

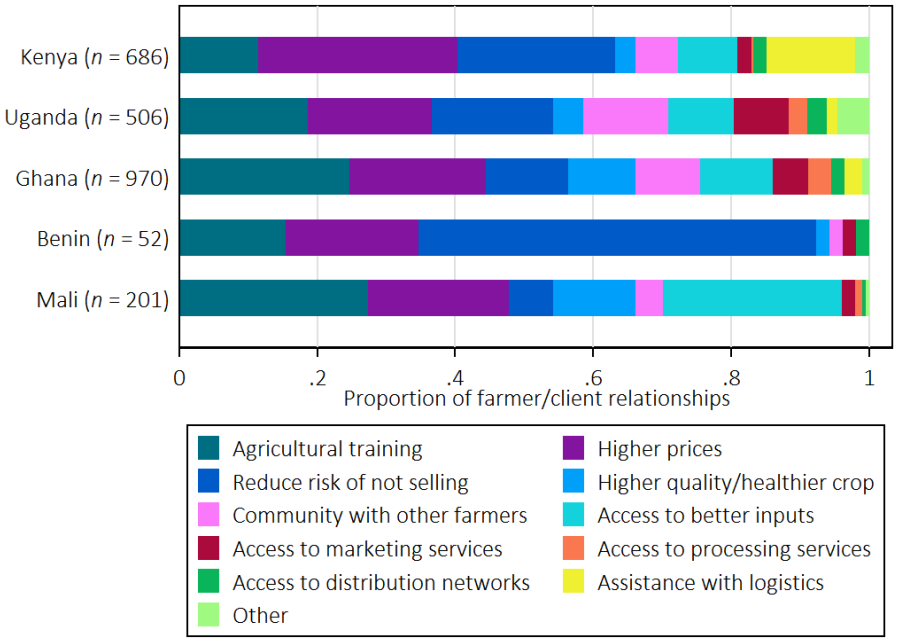
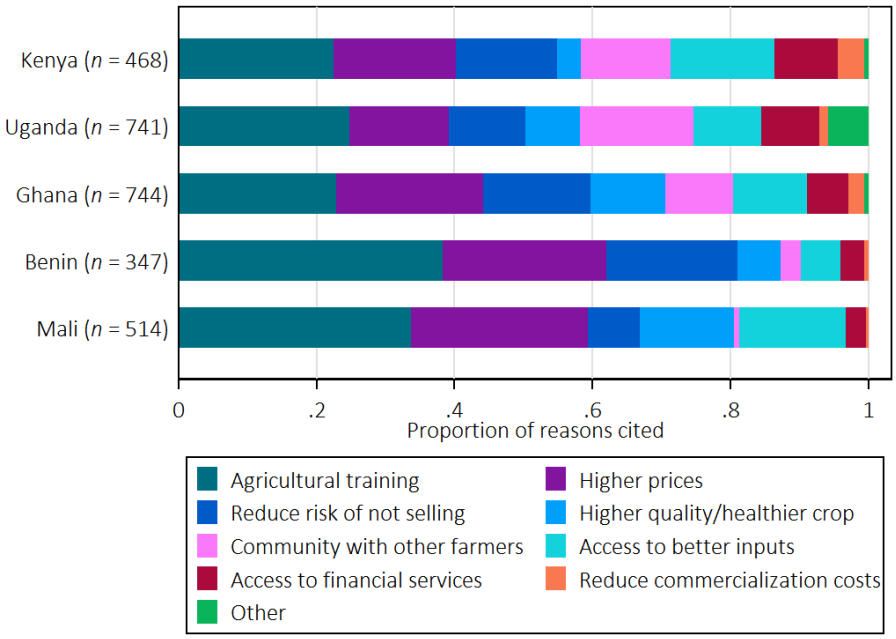


Figure 20. Reasons for Belonging to Producer Group



In all countries, more than 70% of the farmers selling through non-independent product commercialisation methods reported being willing to sell to clients again in the future (Figure 21). Some of the main reasons farmers who said they would sell to clients again in the future were to receive agricultural training, receive higher prices, reduce the risk of not selling, produce a higher quality or healthier crop, and achieve community with other farmers (Figure 22). On the other hand, farmers who said they would not sell to clients in the future commonly reported that they would not sell to get higher prices (Figure 23).

Figure 21. Interest in Selling to Clients in the Future

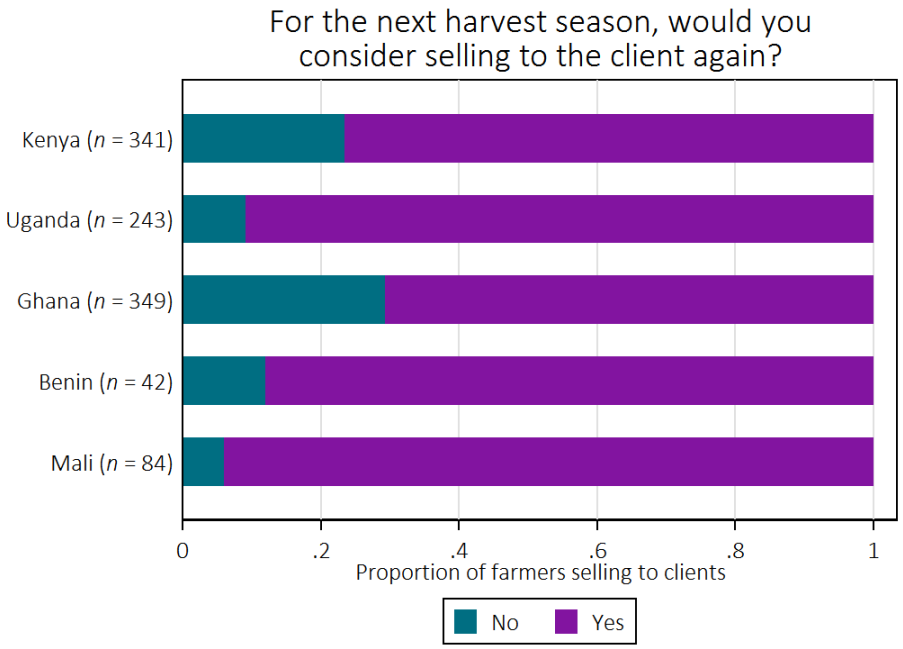


Figure 22. Reason for Selling to Clients in the Future

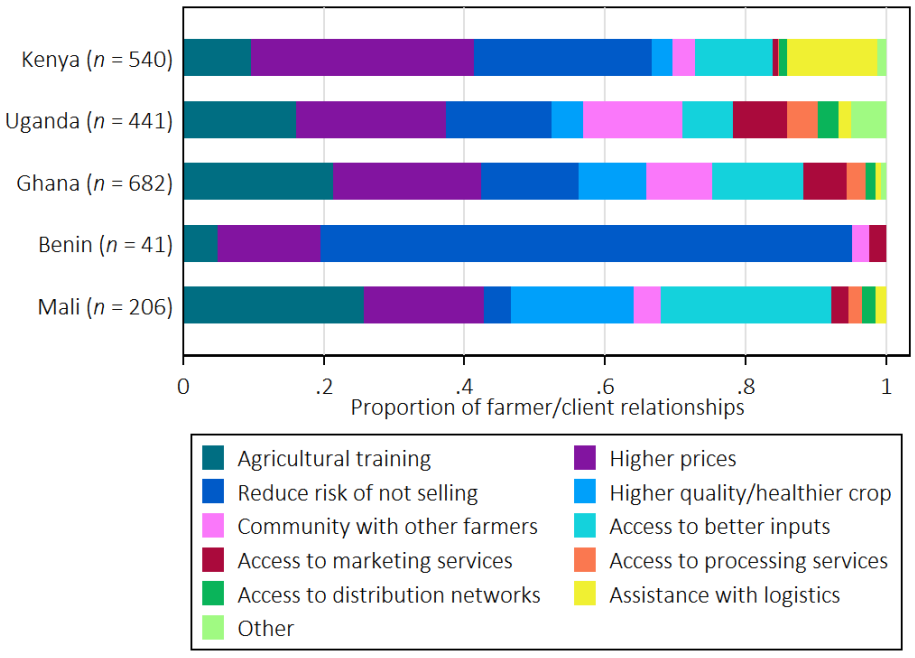
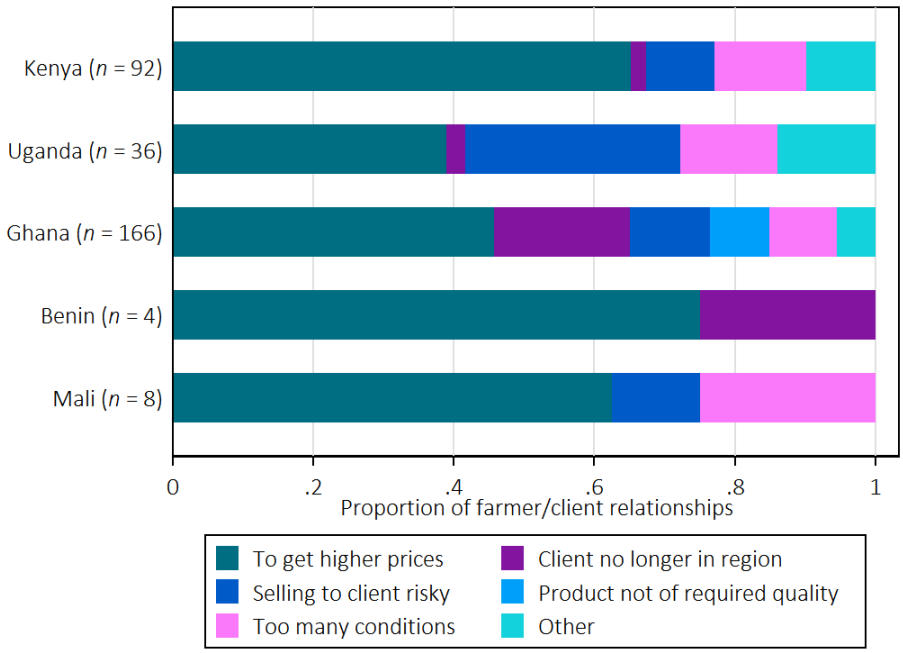


Figure 23. Reason for Not Selling to Clients in the Future



6. Conclusions

The purpose of the Endline Report is to present the results of the final impact evaluation of the 2SCALE programme. The impact estimates used baseline data collected in 2015 and endline data collected in 2017 from five countries where 2SCALE currently operates.

Empirical strategy. To conduct a valid assessment of the impact of 2SCALE on farmers' income and food security, it was necessary to establish a clear counterfactual. This required rigorous methodologies that enabled us to address the question of "What would have happened in the absence of the intervention"? For the evaluation of 2SCALE, we used a Difference-in-Differences (DiD) empirical strategy, which compares changes in outcomes over time between programme beneficiaries and nonbeneficiaries. DiD entails calculating the change in an outcome, such as income, between the baseline and the follow-up period for treatment and comparison groups and comparing the magnitude of those changes. The key assumption underpinning the DiD is that no systematic, unobserved, time-varying difference exists between the treatment and comparison groups. As long as this assumption is satisfied, DiD accounts for potential self-selection from unobserved time-invariant characteristics, which is especially relevant in this evaluation because farmers may self-select into the programme based on characteristics that are not readily available or observable, such as farmers' motivation, ability, or ambition. In other words, the DiD strategy provides a robust methodology to estimate the impacts of the programme, while addressing some potential concerns of unobserved farmer characteristics driving the results. The DiD strategy was conducted in all countries except for Ghana. For Ghana, a country for which no pure baseline exists because the programme had been in operation

across both treatment and comparison groups since 2013, we used a different identification strategy that combines regression and propensity score matching methods on the endline data.

Limitations. The present study has some limitations that are worth discussing. First, for our empirical strategy to produce unbiased estimates of programme impacts, we need that, in the absence of treatment, the difference between the treatment and control group is constant over time. This requirement, known as the parallel trend assumption, can be tested as long as there are at least two rounds of data before programme implementation. With only one round of data before the programme, as is our case, the parallel trend assumption is a non-testable condition.

As shown in the report, our treatment and control groups were purposively located in different areas in order to avoid potential programme contamination effects from the treatment to the control group. While we worked closely with the partnerships and 2SCALE representatives to select control farmers that were as similar as possible to treatment farmers, the results for some outcomes in given countries show large differences in terms of the average characteristics of farmers from both groups. Although the parallel trend assumption does not require that the characteristics of the treatment and control groups have similar means at baseline, the more dissimilar the treatment and control groups are at baseline, the more likely the two groups follow a different trend over time in absence of the programme, violating our key identification assumption. For example, if the treatment and control groups are in very different areas, they may face different agro-ecological and climatic characteristics such as temperature, rainfall, sunshine, or draught. Such differences could vary over time and between areas and bias our estimates. Unfortunately, we don't have auxiliary or anecdotal information (for example from qualitative data collection) that can be used to assess to what extent the observed differences at baseline represent an important limitation to our identification strategy.

A second limitation of the study is that the time frame used for the evaluation may be too short to detect programme impacts on longer-run outcomes such as income, productivity, and food security.

Results. We find some positive evidence that 2SCALE improved outcomes for farmers, especially for some key intermediate outcomes. In Kenya we find statistically significant increases in the use of cultural practices by treatment farmers, which could be linked to the training received from the farmer field schools. In Benin, we find that beneficiary farmers were more likely to adopt some positive practices presumably as a result of the trainings received. In Benin and Uganda, we see an increase in the use of some positive cultural practices and a decrease in farm-level challenges; specifically, we saw a decrease in the likelihood that a farm was affected by weeds, pests or diseases. In Mali, the results show some mixed results with farmers adopting some positive cultural practices, while being subject to a higher likelihood of being affected by insects, fungus, or diseases.

Regarding target crop outcomes, the greatest success is the partnership with Shalem in Kenya, where statistically significant evidence indicated that farmers' sorghum production increased. Little evidence for statistically significant changes by 2SCALE farmers in key production outcomes (quantity of the total harvest, value of the total harvest, or gross margins) was shown in Uganda, Benin, or Mali. In Ghana there were statistically significant decreases in the quantity

and value of the total harvest and the gross margins. However, the results for Ghana need to be interpret with caution because the empirical strategy used for this country does not allow us to control for unobserved determinants that may be correlated with programme outcomes due to the lack of a true baseline.

Moreover, the lack of positive impacts on production and productivity outcomes for most countries (except for Kenya) is not surprising given that two years is a relatively short period for the programme to start producing effects. Also, some of the partnerships had some implementation delays, which may have also affected the likelihood to produce longer-run impacts. In Uganda, our examination of the mechanism through which 2SCALE creates impacts suggests that a key activity was not being provided, which could explain the null findings. That is, a small share of treatment farmers reported that their producer group facilitated access to machinery (Figure 15) and in fact treatment farmers were less likely than comparison farmers to report that their producer group facilitated access to machinery. These findings correspond with information provided by 2SCALE representatives that the oil machinery was only recently established. In Benin, the decreased use of inputs and lack of change in production values could be due to the collapse of the Nigerian market in 2016, stemming from the devaluation of the Nigerian currency, since we understand that many of the treatment farmers supplied to the Nigerian market. In Mali, the lack of findings related to production of the target crop may be due to the nature of the beneficiary farmers included in the study. The farmers in the treatment group in Mali came from two of the larger producer organizations (PO) that we understand happened to have difficulty reimbursing the credit for inputs obtained and also experienced governance challenges. The lack of findings in Benin and Mali could also be due to the fact that a surprisingly large fraction of treatment farmers reported selling all the crop independently in those countries (Figure 7), which suggests that established relationships with clients is important for programme success. Despite the mixed results, there is positive evidence that 2SCALE increases income and food security through the various activities the programme supports.

In the case of Uganda, our examination of the mechanism through which 2SCALE creates impacts suggests that a key activity was not being provided, which could explain the null findings. That is, a small share of treatment farmers reported that their producer group facilitated access to machinery (Figure 15) and in fact treatment farmers were less likely than comparison farmers to report that their producer group facilitated access to machinery. These findings correspond with information provided by 2SCALE representatives that the oil machinery was only recently established. In Benin, the decreased use of inputs and lack of change in production values could be due to the collapse of the Nigerian market in 2016, stemming from the devaluation of the Nigerian currency, since we understand that many of the treatment farmers supplied to the Nigerian market. In Mali, the lack of findings related to production of the target crop may be due to the nature of the beneficiary farmers included in the study. The farmers in the treatment group in Mali came from two of the larger producer organizations (PO) that we understand happened to have difficulty reimbursing the credit for inputs obtained and also experienced governance challenges. The lack of findings in Benin and Mali could also be due to the fact that a surprisingly large fraction of treatment farmers reported selling all the crop independently in those countries (Figure 7), which suggests that established relationships with clients is important for programme success. Despite the mixed results, there is positive evidence that 2SCALE increases income and food security through the various activities the programme

supports. In particular, the partnership with Shalem in Kenya is very promising, a programme that seems very well managed and implemented from the start and is showing very promising results.

We find mixed evidence of the changes in food security of households. Across most of the countries, we find that some subjective indicators of food security were positive. In examining the indices of the food insecurity scale and the household dietary diversity score (which aggregates consumption across food groups), a decrease in the food insecurity scale occurred in Kenya, suggesting that farmers improved their food security, but no statistically significant change occurred in the household dietary diversity score. In Uganda and Ghana, neither the food insecurity scale nor the household dietary diversity score had statistically significant changes. In Benin and Mali, we find evidence of a reduction in food security since we found increases in the food insecurity scales. The household dietary diversity scores in Benin and Mali did not have statistically significant changes. Again, these mixed results on food security and diet diversity outcomes are not surprising given that we are not finding statistically significant effects on income for most countries. Moreover, even programmes that produce positive impacts on income may not show positive effects of food security and diversity because consumption patterns may take some time to be modified. Also, 2SCALE in general does not implement activities to help farmers change their behaviours regarding diets and nutritional aspects of their food intake.

In examining the mechanisms through which 2SCALE operates, we saw some mixed results on outcomes related to non-independent product commercialisation. The number of clients to which treatment farmers sold decreased in Benin (possibly due to the collapse of the Nigerian market) and was not statistically significant in Mali. Treatment farmers in Kenya were more likely to report that they would sell to their clients again in the future. The opposite was the case in Benin and the results for this question were not statistically significant in Mali. In terms of what aspects of production clients facilitated access to, the countries varied both in the specific aspects reported and whether those aspects had a positive or negative impact.

Regarding the impact across network or social capital dimensions, the impact estimates were mixed in Uganda regarding aspects of production to which producer groups facilitated access. Treatment farmers in Uganda were more likely to report that their group facilitated access to credit and processing, activities that the partnership provided. Most likely due to the delay in establishing the oil mill, treatment farmers were less likely to report that their group facilitated access to machinery. The impact estimates in Ghana were more favourable, in that there was an increased likelihood that treatment farmers' producer groups facilitated access to seeds, fertiliser, machinery, professional sprayers, credit, and marketing. However, the opposite effect was documented in Benin in that the treatment groups in those countries reported a decrease in the likelihood of access to production aspects. The results for these measures were not statistically significant for Mali, and producer groups were not relevant in Kenya.

In our graphical examination of awareness, use, and preferences, treatment farmers reported that clients and producer organisations offered a diverse set of activities and inputs, but the list of components that farmers reported using was more concentrated. Nevertheless, across all countries, more than 70% of farmers selling to clients reported being willing to sell to clients again in the future. Some of the main reasons farmers would choose to sell to clients in the future

were to receive agricultural training, receive higher prices, reduce risk of not selling, produce a higher quality or healthier crop, and achieve community with other farmers. On the other hand, farmers who said they would not sell to clients in the future commonly reported that they would not sell to get higher prices, which corresponds to the fact that low prices or price fluctuations seemed to be the more commonly reported problems in selling to clients.

Overall, the results show that some of the 2SCALE partnerships assessed have the potential to change the behaviour of farmers (intermediate outcomes). Some of the intermediate outcomes analysed as well anecdotal information from programme implementers show that 2SCALE is improving partnership governance, business models of the lead firms (when applicable), access to finance, timely payment, adoption of innovations, access to markets, among other things. If these changes in intermediate outcomes are sustained over time, the different partnerships may experience longer-run impacts on key final outcomes such as yields and income.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies, 72*(1), 1–19.

Bellemare, M. (2012). As you sow, so shall you reap: The welfare impacts of contract farming. *World Development, 40*(7), 1418–1434.

Bonilla, J., & Rai, N. (2015). *Evaluation of 2SCALE: Inception report.* Washington, DC: American Institutes for Research.

Bonilla, J., & Rai, N. (2016). *Evaluation of 2SCALE: Baseline report.* Washington, DC: American Institutes for Research.

Cavatassi, R., Gonzales, M., Winters, P., Andrade-Piedra, J., Espinosa, P., & Thiele, G. (2011). Linking smallholders to the new agricultural economy: The case of the *Plataformas de Concertación* in Ecuador. *Journal of Development Studies, 47*(10), 1545–1573.

Coates, J., Swindale, A., & Bilinsky, P. (2007). *Household Food Insecurity Access Scale (HFIAS) for measurement of food access: Indicator guide* (v. 3). Washington, DC: Food and Nutrition Technical Assistance Project. Retrieved from https://www.fantaproject.org/sites/default/files/resources/HFIAS_ENG_v3_Aug07.pdf

Demirguc-Kunt et al., 2015, *Global Financial Inclusion Database*, World Bank.

Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica, 66*(5), 1017–1098.

Heckman, J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. *Review of Economic Studies, 64*(4), 605–654.

Imbens, G., & Wooldridge, J. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature, 47*(1), 5–86.

IRCA. (2015). *Toward Sustainable Clusters in Agribusiness through Learning in Entrepreneurship (2SCALE): Strategic work plan 2015–2017.* Wageningen, The Netherlands: Author.

Kidoido, M. M., & Child, K. (2014). *Evaluating value chain interventions: A review of recent evidence* (ILRI Discussion Paper 26). Nairobi, Kenya: International Livestock Research Institute.

Pradhan, M., & Ravallion, M. (2000). Measuring poverty using qualitative perceptions of consumption adequacy. *Review of Economics and Statistics, 82*(3), 462–471.

Robins, J., & Rotnitzky, A. (1995). Semiparametric Efficiency in Multivariate Regression Models with Missing Data. *Journal of the American Statistical Association,* 90(429), 122–29.

Robins, J., Rotnitzky, A., & Zhao, L. (1995). Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data. *Journal of the American Statistical Association*, 90(429), 106–21.

Roduner, D. (2005). Value-chains: What is behind this "new" key word? And what is the role of development agencies in value chain development in developing countries. *Rural Development News, 2,* 10–15.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inferences.* Berkeley, CA: Houghton Mifflin.

Smale, M. (Ed.). (2006). *Valuing crop biodiversity: On-farm genetic resources and economic change.* Wallingford, UK: CABI Publishing.

Van der Laan, M., & Robins, J. (2003). *Unified Methods for Censored Longitudinal Data and Causality*. New York: Springer, Physica-Verlag.

Appendix A. Comparison Group Selection Information Sources

To create the sample at baseline, IFDC's country representatives were asked to identify areas from which approximately 400 treatment and 400 comparison farmers could be enumerated. Either through discussions (in the case of Kenya) or email exchanges we worked with IFDC representatives to select the control area. For each of the countries, we asked key individuals8 to fill out a brief questionnaire, designed to give us an idea of the background of the program, where it was currently operating, where it planned to expand, and where it was not working. We wanted to know about areas that were not too near the treatment areas, and we wanted to know about areas where the partnership was not planning to work within the next two years to prevent spillovers (where the program benefits spill over to the comparison group) from biasing the results. However, the areas also needed to have a substantial number of farmers of the target crop to ensure that we would capture an appropriate sample size. Answers to these questions informed our selection of the control area.

Kenya: We had discussions with Ruth Kinoti N'ee Mbogori, who helped identify control areas based on the following criterion:

- The areas where Shalem isn't operating in now.
- Clusters where they would operate in if they had the capacity.
- Clusters which are far away from current treatment clusters (in fact they did identify clusters in another County (Tharaka Nithi County)
- Clusters where they are likely not to go to in the next two years.

Uganda: We worked with Amos Kisilu who said "We have identified three cotton growing areas; Kyenjonjo, Hoima and Masindi Districts. We propose Masindi District for the control group because Allan Wayira is very familiar with Masindi District. The area has active cotton growing farmers. However, caution should be taken not to select farmers receiving support from other projects."

Ghana: Ernest Acheampong's responses to our emails identified "Several communities in Wapuli sub-district of the Saboba district" as areas with a large amount of soybean farmers who don't participate in 2SCALE. Gabriel Mills wrote; ".... As I explained earlier 2SCALE has no intention of rolling out any intervention in Yendi. Again, our clusters are within a specific geographical location. The closest cluster to Yendi is Saboba hence if the control group are located away from the communities on the Yendi –

[8] We contacted Amos Kisilu (then 2SCALE staff), Ruth Kinoti N'ee Mbogori the partner facilitator in Kenya; Amos Kisilu, Joseph Mwaka (then 2SCALE staff) and Allan Wayira (then 2SCALE staff) in Uganda; Ernest Acheampong and Gabriel Mills in Ghana; Tonato Oliver, C Addupong, Eric Lakoussan, and Ernest Acheampong in Benin; and Frederic Sanogo and Ernest Cheampong in Mali. We also liaised closely with Jan Williem Van Casteren (then 2SCALE staff).

Saboba road that will be fine. That is not to say that all communities on the Yendi-Saboba road are part of the Saboba cluster...”

Benin: Olivier Tonato and C Addupong responded to our questions. Eric Lakoussan also responded and wrote, “As group control, we can use the producer from Aplahoué. They are more than 400 and they produce both pepper and tomato.”

Mali: We got responses back from Frederic Sanogo to the question: Are you able to identify some areas with a large number of maize farmers who don’t supply to SONAF? If yes, can you please provide the names of these areas? He provided the following; Woroni (commune of Loulouni), Fanidiama, Nassoulou (commune of Zégoua), Zérelani (commune of Kléla), Foh (commune of Kourouma) and the rural commune of Lobougoula

Appendix B. Identification Strategy for Ghana

Because no pure baseline existed for Ghana, we estimated the impact results using an identification strategy that relied only on the endline dataset. The identification strategy proposed to estimate the causal effects of the programme relies on the doubly robust estimator developed by Robins and Rotnitzky (1995); Robins, Rotnitzky, and Lue Ping Zhao (1995); and van der Laan and Robins (2003). At the farmer level, the proposed approach combines regression and propensity score matching methods in a three-step approach to estimate treatment effects.

In the first step, a treatment model is defined that explains the probability of programme participation. Specifically, we estimate the probability of programme participation through a logit or probit such as:

$$T_i = \Phi(\boldsymbol{X}_i \cdot \boldsymbol{\beta}' + \varepsilon_i)$$

where T_i is a dummy for having received the programme, X_i is a vector of individual, and ε_i is an error term. The observable characteristics considered in the treatment equation controls for household size at baseline, dummies for household age composition; main respondent's gender, age, marital status, years of farming experience, education, use of computer, internet, and mobile phone; distance to centre of town; dummies for most important parcel being black/dark coloured, very fertile, flat or slightly flat, and has no erosion; number of rooms; dwelling is owner-occupied; shelter is traditional house; dummies for type of walls; dummies for type of floors; dummies for types of water access; household has electricity; dummies for types of toilet; and asset index. Then we generate propensity scores,P_i, the probability of receiving treatment, as

$$P_i = \Phi(\boldsymbol{X}_i \cdot \widehat{\boldsymbol{\beta}}')$$

From the first step, inverse-probability weights are derived from the estimated propensity score.

Second, using the estimated inverse-probability weights, weighted regression models are fit for the outcome equation for each treatment level and obtained the treatment-specific predicted outcomes for each subject. Lastly, means of the treatment-specific predicted outcomes are computed and the difference of these averages provides the estimate of the average treatment effect of the programme. Intuitively, weighting can be interpreted as removing the correlation between the treatment condition and other covariates that may be correlated with treatment, and regression as removing the direct effect of such variables on the outcomes of interest (Imbens & Wooldridge, 2009).

This approach assumes that programme participation is exogenous to potential outcomes conditional on observable characteristics—that is, that there is no selection bias due to unobserved characteristics and that the observable characteristics we capture determine programme participation. Due to the unobservable nature of these potential additional characteristics, this assumption is untestable. Nevertheless, we employ a series of strategies to reduce the potential threat of the impact estimates being driven by unobserved characteristics of programme participants. Specifically, we used a filter questionnaire to replicate the selection of potentially eligible farmers, and to collect numerous covariates as controls that are good predictors of programme participation. Several authors have argued that social programmes can be evaluated using matching methods, if there is access to a rich set of variables that determine programme participation, and that the nonexperimental comparison group is drawn from the

same local region as participants (Heckman, Ichimura, & Todd, 1997; Heckman, Ichimura, Smith, & Todd, 1998).

Appendix C. Fieldwork Report for Endline Data Collection for the Impact Evaluation of IFDC-2SCALE Farmers Survey 2017

Dalberg Research

FIELDWORK REPORT

FOR

Endline Data Collection for the Impact Evaluation of IFDC-2SCALE Farmers Survey 2017

Dalberg Research

Research Solutions Africa Ltd (RSA) becomes **Dalberg Research**, a company of the Dalberg Group. Teaming up with Dalberg means that we are now able to offer a wider range of products and services. Dalberg's demanding standards also put us on a faster track of growth in terms of client experience, quality of delivery and depth of insights. For those who have worked with RSA before, very little changes. You will find the same management and the same growing team of researchers at the same office location. All contracts and confidentiality agreements (formal and informal) remain intact. Without your explicit consent, access to any confidential and proprietary material shared is limited to Dalberg Research. If you have any questions on this development, kindly reach out. The current legal name of the company is Research Solutions Africa Ltd, and Dalberg Research is a brand name under which Research Solutions Africa Ltd operates. The legal name is expected to be aligned with the brand name in the coming months.

The Endline Data Collection for the Impact Evaluation of IFDC-2SCALE Farmers Survey 2018

Table of Contents

Executive summary

This report gives an overview of how endline data collection and related activities were implemented ***the Endline Data Collection for the Impact Evaluation of IFDC-2SCALE Farmers Survey 2018***, in the period between September and December 2018 across 5 countries *Kenya, Uganda, Mali, Benin and Ghana*. It outlines the methodology used during the data collection, the teams involved, the challenges encountered and how these were mitigated.

During the implementation of the fieldwork, in each country, we worked three levels of staff; the enumerators, Team Leaders and Supervisors of different sizes they were identified, trained and engaged in the survey. Overall fieldwork supervision was provided by Dalberg Research Field Coordinator and the Project Manager.

By the end of the exercise a total of 1931 farmer household were interviewed as treatment respondents and a total of 1909 farmer household were interviewed as control respondents, with an achieved output of 3840 out of the targeted 4002 interviews.

1. Project Background

IFDC 2SCALE Project is funded by the Dutch government through the Directorate General for International Cooperation (DGIS, The Netherlands) from 2012 to 2017. 2SCALE develops a portfolio of agribusiness clusters in target countries in sub-Saharan Africa (Benin, Ghana, Mali, Ethiopia, Kenya, Mozambique, South Sudan, Uganda, Cote D'Ivore and Nigeria). These clusters are defined as a partnership at the local level among actors (producers and their organizations, input suppliers, finance suppliers, processors, warehouses managers, traders, business development services, etc.) that share a common ambition to build profitable commodity-based value chains. The agribusiness cluster approach is designed to help rural smallholders move from subsistence farming to fanning as a business and supply agricultural products for local, national, regional, and international markets. The project partners with national and multinational agri-food enterprises, as drivers, to increase productivity and to improve efficiency and sustainability of supported agribusiness clusters

1.1. The survey Objective

To be able to ascertain the progress of program implementation and its impacts of 2SCALE, a Performance Monitoring Plan (PMP) was designed to guide the acquisition of timely and relevant information for project management, partners and donors. This enabled management, donors and partners to gain feedback and deduce lessons, determine the direction and magnitude of progress, and make adjustments to interventions for effective translation of project inputs into outcomes.

As part of the operationalization of the PMP, surveys were designed to be conducted in the selected countries. The baseline and endline were designed to include; first establishing baselines and secondly providing periodical data. The data was collected as per the cluster (farmer) level indicators in the PMP.

Farm-household surveys were conducted to evaluate the impact of 2SCALE interventions, focusing on level of income as well as food and nutrition security of participating farmers, across various product groups and partnership types.

Farm-household surveys were carried out to evaluate the impact of its interventions, focusing on level of income as well as food and nutrition security of participating farmers, across various product groups and partnership types. The IFDC/2SCALE endline survey was carried out in Kenya, Uganda, Mali, Benin and Ghana. Ethiopia was excluded from the endline due to an ongoing embargo at the time of the survey.

2. Set up activities

The formulation of the survey questionnaire to be used in the endline survey was undertaken by AIR. Dalberg Research reviewed the questionnaire and undertook translation activities into the relevant languages in each country.

Under the set-up activities, Dalberg Research undertook the following activities in each country; programming of the survey tool, enumerator identification and training, piloting, pilot debrief, selection of the final field team members, and deployment of the same to the various regions in the survey countries. The training reports for each country outlines activities undertaken during enumerator identification and training, piloting, pilot debrief, selection of the final field team member.

This report outlines the field activities in the five countries Kenya, Uganda, Mali, Benin and Ghana.

2.1. Programming of the finalized questionnaire into the HHDs

To allow for data capture through CAPI, we programmed both the English and local versions of the questionnaire using 'Dooblo Survey to Go' software. The programming was done internally, by one of DR Data Processing staff members. The following languages were used in the respective countries.

Country	Lang: 1	Lang: 2
Kenya	English	Kiswahili
Uganda	English	Luganda
Mali	French	French
Benin	French	French
Ghana	English	Dagbani

To confirm the accuracy and completeness of each of the soft-versions of the questionnaire we implemented several independent mock interviews using the phones to detect and correct as appropriate any likely errors and/or issues like erroneous skips, wrong question numbering or phrasing, incomplete questions, missing questions, etc. that needed to be corrected before the actual fieldwork could start. We did these in very close liaison with the AIR team and the respective 2SCALE team members who attended the trainings in each country. The process ensured that we approve the accuracy, completeness and appropriateness of each and all the questions in the CAPI before fieldwork could commence.

Dalberg Research/AIR and 2SCALE team continuously reviewed the questionnaires (CAPI versions) and the comments were continuously incorporated into the various CAPI versions. With the final approved version, fieldwork commenced each country.

2.2. Sampling

In each country, the sampling frame was made up of farmers who were interviewed in 2015, a list of farmers was generated and shared with respective country teams. Also, generated was Household Information Tracking Sheet which contained contact information for ease of reference by the surveyors in tracking the farmer respondents.

Below is an example of the contact sheets used during the endline survey.

2017 - Household Information Sheet *** COUNTRY: Kenya; LOCATION: Ciothirai

Question	Value in 2015
Household Unique ID	14
Mobile number Line 1	716,555,950
Mobile number Line 2	763,793,791
Mobile number Line 3	
Mobile number Line 4	
Mobile number Line 5	
Mobile number Line 6	
1A. Number of parcels HH OWNED in 2015	3
2A. Number of parcels OWNED used for crop prod.	3
3A. Same as 2A AND at least 1 crop Under Contract (No. parcels)	1
4A. Number of parcels RENTED/SC used for crop prod.	2
5A. Same as 4A AND at least 1 crop Under Contract (No. parcels)	0
6A. Number of parcels HH OWNED or RENTED/SC in 2015	5
1B. Area parcels OWNED in 2015 (acres)	7.0
2B. Area parcels OWNED and used for crop prod. (acres)	1.5
3B. Same as 2B AND at least 1 crop Under Contract (acres)	1.0
4B. Area parcels RENTED/SC and used for crop prod. (acres)	2.0
5B. Same as 4B AND at least 1 crop Under Contract (acres)	0.0
6B. Area parcels OWNED or RENTED/SC in 2015 (acres)	3.5
GPS: Latitude	0.010793
GPS: Longitude	37.806585
Country	Kenya
Location	Ciothirai
Name Main Respondent in 2015	Emma Maina
Local Name	Nyeri
Name HH Head in 2015	Geofrey Magambo
Head's Local Name	Maina
SbjNum 2015	23902877
Respondent Type	Treatment

Crop List in 2015	All Crops
Household Unique ID	14
Crop 1	Local Maize
Crop 2	Sorghum
Crop 3	Beans
Crop 4	
Crop 5	
Crop 6	
Crop 7	
Crop 8	
Crop 9	
Crop 10	

Attached to this report as Annex 3 is the list of HH Information Tracking sheets used for the Endline Survey for each country.

2.3. Recruitment of survey enumerators

Recruitment of survey enumerators was done in each country. In Kenya and Uganda DR recruited the field team members from our internal data base of experienced personnel with whom we have had some working experience in the respective countries. In Ghana, Mali and Benin the fieldwork partners did the same.

In identifying the candidates for the training and eventual engagement in the data collection activities in the survey for each of the 5 countries, we were guided by five key qualities: individual fieldwork experience; level of education; past performance record in assignments in specific the 2scale 2015 baseline survey and availability during the entire data collection phase. Details of enumerator recruitment for each country is contained in the country training reports.

2.4. Enumerator training

Training of the enumerators was undertaken in each country with country specific questionnaires. During the trainings we discussed in detail the general background of the survey; why the survey was being undertaken in the identified target areas; survey objectives; the survey questionnaire (both the French/English and local versions as appropriate); use of the phones in administering the interviews; the survey samples (target respondents, sample points and sizes) and the related sampling approaches; the survey timelines; the standard fieldwork procedures to be implemented during fieldwork, including data quality control measures and ethical issues applicable; client expectation of the data collection team; modalities for handling field challenges and related issues; logistical plan in the survey; payment terms and the applicable contracts; and the communication protocol to be used during the survey.

The survey questionnaire review entailed reading and discussion of each of the questions, and general paired & plenary mocking sessions by the participants to assess the flow, consistency and appropriateness of the phrases and terms used therein. We used both the paper- and phone-based versions of the questionnaire during the training. The same procedure was done in every country.

2.5. Piloting

After the trainings, the participants in each country were taken through a one-day piloting session to assess their understanding of the survey tool and the related field procedures as outlined during the training. The exercise was also geared at checking how effective the survey tool was in capturing the various responses from the survey respondents, especially with reference to the flow, consistency and appropriateness of terms and phrases used in the questions.

After the country specific piloting exercise, debriefs were done to reflect on all the lessons learnt.

2.6. Pilot de-brief

Pilot debrief sessions were geared at capturing the participants' experience in the field during the pilot exercise: how effective the survey tool was in bringing out appropriate responses from the survey respondents, any questions needing further clarifications or re-phrasing, any challenges encountered with the tool or survey respondents, and any recommendations on how to address any of the issues arising out of the pilot. The present 2SCAL team and the Project Manager responded and gave appropriate survey-perspectives to each of the questions raised by the participants with respect to the survey questionnaire/s.

The observations were reported by the participants during the de-briefs: the same is contained in each country training report.

3. Fieldwork

In terms of personnel levels, in each country, we had enumerators, team leaders and supervisors, each with specific duties and responsibilities to undertake, for an effective implementation of the data collection activities in the survey. The responsibilities were similar across the countries.

3.1. The Enumerators responsibilities

The enumerators were responsible for the actual administration of the face to face interviews to the respective target household respondents, using the Huawei IDEOS and Samsung smart phones in Kenya, Uganda, Mali and Benin. In Ghana data collection was done using tablets. They worked under close leadership of the team leaders.

3.2. The Team Leaders responsibilities

The team leaders were responsible for ensuring that the quotas set for their sub-teams were accomplished efficiently, rightly and using the recommended quality control procedures. They assigned specific interviews to their respective enumerators, ensured that the study respondents were correctly sampled, and that the interviews were administered as recommended. They undertook quality control checks on the enumerators assigned to them in the field through regular random sit-ins, call backs and back checks.

The team leaders liaised very closely with the respective team supervisors and or the survey coordinator as appropriate in addressing or reporting on any noted challenges in and fieldwork progress by their respective sub-teams.

They were responsible for paying the relevant courtesy calls to the respective cooperative presidents in each of the target cooperatives as appropriate, and in ensuring that there was a cordial working relationship amongst the various sub-team members. They reported directly to the Field Supervisors and or the Survey Coordinator, who then relayed any pertinent issues to the Project Manager and/or DP Manager.

3.3. Field Supervisors

The field supervisors coordinated all the data collection activities by the enumerators and team leaders for the respective teams. They reported on their respective teams' daily progress (challenges encountered, and outputs achieved), and also undertook similar random quality control checks as the team leaders on the enumerators.

4. Data collection

Fieldwork was implemented between September and December 2017, at the Respondent household level, the enumerators used purposive sampling approach to identify the correct adult household member to engage in the survey, this was determined during the baseline survey in 2015. We interviewed the household member most knowledgeable about farming activities. Again, we did the same during the endline data collection,

The mode of data collection used in the survey in each country was mainly phone-based face to face interviews using the 'Dooblo survey to Go' software to record and transmit the data.

By the end of the data collection exercise, we had effectively administered a total of 3,840 (Three thousand four hundred farmers) successful interviews, out of the targeted 4002. The distribution of the countries complete surveys is summarized in Table 1 below:

Table 1: Achieved fieldwork output per country

	Target	Treatment	Control	Total Achieved
Kenya	800	377	379	756
Uganda	800	362	386	748
Mali	800	400	400	800
Ghana	802	406	370	776
Benin	800	386	374	760
TOTAL		1931	1909	3840

4.1. Field supervision and data quality assurance

In each country, each sub-team had a Team Leader who was mainly responsible for the co-ordination of field work. All Team Leaders undertook regular quality control checks with their enumerators to ensure that the data collected were accurate, complete, and from the right respondents who were identified and selected using the recommended sampling procedures for the survey. They used sit-ins, back checks and call backs to confirm and ensure that data quality control procedures were being observed by the enumerators as planned.

Above the Team Leaders were the Field Supervisors whose main responsibilities were to co-ordinate all the activities by the enumerators and team leaders assigned to them, and capture and report as appropriate on any challenges encountered by their teams and the related daily outputs. They undertook similar quality control checks as the team leaders on the activities by the enumerators and reported to the Field Supervisor for Dalberg Research in Kenya, PSI Field Supervisor in Ghana and Field Supervisor for DADACH for Mali and Benin.

The DR Field Coordinator liaised with the Data Processing Unit to ensure that the outputs reported in the field were the same as those reported from the survey server. He also followed up with the given enumerators and team leaders on all cases flagged as possible errors from the data quality control checks implemented on the results of the daily server downloads from the field.

4.2. Survey Challenges and How They Were Mitigated

Overall, majority of the survey respondents were quite willing to participate in the survey. However, we did face a number of challenges while implementing the data collection activities in the survey, herein captured under four main sub-types: location-related challenges; respondent-related challenges; technical challenges, enumerator-related challenges and other challenges/issues related to fieldwork

We describe each of these challenges below:

4.2.1. Location-related challenges

Lack of power to charge the phones, especially in the rural areas

In some survey areas in different countries the enumerators had a challenge in charging the phones and power banks used for the data capture since there was no electricity in the area or there was power rationing or random power blackouts in the area on certain days.

***Action**:*

- *The affected teams ensured that they rushed to their rooms and plugged the phones and power banks to charge before they sourced for food, etc. for the evening; this was to allow the phones to be fully charged by the time power went off. In other areas they made sure that the phones and the power banks were fully charged before venturing into given areas so that they could be able to effectively undertake all the interviews before the phones ran out of charge.*
- *The enumerators were reminded to conserve the power charge of the phones by not accessing the internet using the survey phones or having the GPS function on throughout the day; they were to put it on only when they were actually administering the household interviews.*

4.2.2. Respondent-related challenges

Some respondents did not want to take part in the survey again.

Farmers mainly from control areas were reluctant to be interviewed they said that they had been previously interviewed and did not benefit in anyway, while they knew of some farmers who had benefitted in some programs. We were not able to establish whether this was in direct relation to other farmers benefiting from 2SCALE related programs or other programs.

***Action**: Our fieldwork partners informed the presidents and some interviews were eventually conducted though some were not conducted a report of the same is as indicated in the attrition cases in every country.*

Farmers who were no longer working with the appointed 2 SCALE aggregators.

For example, in Kenya there were farmers who were no longer working with SHALEM.

***Action**:*

- *For such cases we sought the farmers consent on being interviewed and they did accept to be interviewed, we included them in the sample. There was a question introduced for each country whether or not the farmer was still working with the aggregator they worked with in 2015.*
- Some farmers though completely refused to be interviewed indicating they no longer worked with Shalem

In Benin control area; 10 respondents started the interview and stopped midway: This was unusual, since we did not have any experience like this in any other country. We also established that it was not unique to one/few enumerators.

4.2.3. Technical challenges

Completed interviews not reflected in the survey server - Benin
Some 7 interviews which had been done and assumed to be synced did not reflect in the server, during the final reconciliation of completed interviews as reported by the field team and as reflected in the server. The concerned enumerators re-visited those particular respondents and explained the anomaly which was probably caused by phones hanging and redid the interviews.

Delayed uploading of completed interviews – Ghana/Uganda/Mali/Benin
Some teams were unable to load interviews on a daily basis because of network issues, however, they were able to do so once the respective teams in respective countries were in an area with network connection.

GPS capture issues – Benin/Ghana

In a few cases, the phones used could not promptly capture the GPS coordinates for the households where the interviews were administered.

Action*:* The enumerators had been informed on how to handle this challenge; they made several repeat-capture trials, until the coordinates were captured.

4.2.4. Enumerator-related challenges

Erroneous capture of given PSU names by some enumerators
Some enumerators erroneously selected wrong names for given cooperatives a problem attributable to lack of keenness on the enumerators' part while selecting the appropriate cooperatives names in the phones.

Duplicate IDs
From the error report run on uploaded data it was apparent that the enumerators were assigning duplicate IDs
Action**:** Enumerators were asked to clean up the IDs. The importance of using notebooks was stressed.

In addition to the above challenges, listed below are other survey challenges which were country specific and explanations on the attrition cases per country;

4.2.5. Field challenges - Kenya

- There were instances where the Farmers who were interviewed during baseline are no longer planting sorghum and for that reason they said they were not willing to be interviewed
- Some farmers no longer work with Shalem as the aggregator and for that reason they were not willing to be interviewed.
- There were households who migrated to different towns and don't farm anymore
- There were farmers who had disagreements with Shalem and refused to be interviewed
- There were farmers who travelled out of town and no one knew when they would be back
  Other challenges which had impact on time included;

- The respondents who had been away for a long time and there would be a possibility of them coming back within the week, our team revisited these farmers and conducted interviews.
- The respondents whose phone numbers weren't going through for some days but were able to come online and appointments booked with them and interviews eventually conducted.
- Respondents who weren't known but got identified through further snowballing.

4.2.5.1. Attrition cases - Kenya

These cases included;

	Name	Village	HH ID	Treatment/Control	Total
There were instances where the Farmers who were interviewed during baseline are no longer planting sorghum and for that reason they were not willing t be	Rose Gitonga	Kianjogu	146	Treatment	4
	Paul Anaboi	Mbirikene	382	Treatment	
	Japheth Mbundi	Mbirikene	358	Treatment	
	Kaai Mitambo	Kamujwa	786	Control	
The team also came across farmers we interviewed but are no longer no longer working with Shalem as the aggregator, and for that reason they were not willing to be interviwed	Joseph Mwiti Kiria	Gachua	130	Treatment	3
	Stella Muthee	Kiangoju	149	Treatment	
	Juster Kirimi	Mukindu	193	Treatment	
There were households who migrated to different towns and don't farm anymore	James Mung'eri	Ciothirai	46	Treatment	13
	Mary Andrew	Kianjogu	154	Treatment	
	Jackline Maitha	Mbirikene	327	Treatment	
	Alex Kimathi	Mukindu	229	Treatment	
	Cecilia Gaite	Mukindu	187	Treatment	
	Raphina Mukiri	Mukindu	192	Treatment	
	Nancy Kathini	Ngaiini	566	Control	
	Jane Wawira	Gaceraka	453	Control	
	Angelina Nzambi	Gaceraka	468	Control	
	Elizabeth Karimi	Nyakinjeru	431	Control	
	Silas Kirimi	Nyakinjeru	436	Control	
	Paul Mutunga	Kamujwa	616	Control	
	Judith Kanunu	Mbirikene	273	Treatment	
There were farmers who had disagreements with Shalem and refused to be interviewed	Mary Kanyore	Mbirikene	253	Treatment	2
	Cosmas Mwenda	Mukindu	246	Treatment	
There were farmers who travelled out of town and no one knew when they would be back	Judith Mukima	Ciothirai	17	Treatment	10
	Elosy Gakii	Mbirikene	288	Treatment	
	Rose Kiambi	Mbirikene	351	Treatment	
	Jacinta Gacheri	Mbirikene	323	Treatment	
	Peter Kirimi	Tunyai	636	Control	
	Geofrey Gathega	Kamujwa	694	Control	
	Ruth Kibare	Gaceraka	481	Control	
	Alice Karea	Ngaiini	581	Control	
	Elizabeth Joseph	Mukindu	232	Treatment	
	Denis Mutethia	Mbirikene	364	Treatment	
There were Farmers Who were interviewed during Baseline but are now Deceased and homesteads Vacant	Zippora Ngomoko	Ciothirai	33	Treatment	1
Farmers Not reachable through phone numbers and Not known in the indicated locations	Maupia Kaaua	Ngaaini	568	Control	11
	Margaret Kawira	Gaceraka	439	Control	
	Francis Muthee	Nyakinjeru	416	Control	
	Josephat Mugambi	Nyakinjeru	427	Control	
	John Mugwe	Nyakinjeru	433	Control	
	Evangeline Mauki	Kamujwa	681	Control	
	Peter Musee	Kamujwa	714	Control	
	Henry Mugendi	Kamujwa	762	Control	
	Rose Gitonga	Kianjogu	161	Treatment	
	Carolyne thuranira	Mbirikene	285	Treatment	
	Paul Anaboi	Mbirikene	382	Treatment	
TOTAL					44

4.2.6. Field challenges - Uganda

The following are some of the reasons we had a high attrition rate in the treatment clusters in Uganda in specific:

- **Kingdom secession**; In Nov 2016 violence broke out between the King's followers and the Ugandan military, the kingdom wanted to form a state comprising of kasese district, the instability led to human displacements which affected some of the survey respondent farmers.
- **Tribal clashes**; Farmers also indicated that in 2016 there was ethnic tension between the Bakojo and the Bamba community which resulted to loss of lives on both sides. This again caused some people to relocate, some to unknown destinations.
- **Drought**: Kasese District was reported to have experienced prolonged drought resulting to total crop failure, this led some farmers to migrate to the mountains across the border to do farming there because the rains were said to be reliable.
- **Cholera Outbreak**; During the survey period there was a cholera outbreak affecting Bwera Katojo, Karambi and Nyamambuka cooperatives. Though there were teams from red cross and UNICEF in the above clusters assisting in containing the outbreak, some deaths occurred. The outbreak also caused disorientation in these three clusters affecting overall achievements

4.2.6.1. Attrition cases – Uganda

	Unique HH ID	Cluster/Co-op	Reason
1	325	Kabirizi coop	Refused to be interviewed due to differences with the cooperative
2	302	Kabirizi coop	Couldn't be traced during the survey and no one knew when he would be back
3	806	Kabirizi coop	Migrated from Kiburara village and couldn't be traced
4	232	Kabirizi coop	Migrated to another locality together with the family
5	277	Kabirizi coop	Migrated with the family
6	243	Kyamihoko	The correct case is 240. In 2015 the respondent in HH ID 243 impersonated Monday Jotham only for the real Monday to show up after the interview. The enumerator had to interview the real Monday hence the duplication.
7	214	Kyamihoko	Refused to be interviewed due to claiming the coop doesn't benefit him in any way
8	530	Rugandabara	Shifted
9	204	Rugandabara	Outright refusal
10	761	Rugandabara	Migrated
11	769	Rugandabara	Migrated
12	665	Rugandabara	Outright refusal
13	14	Rugandabara	Shifted
14	472	Bwera Katojo	Outright refusal

15	771	Bwera Katojo	The respondent died and there was no else to be interviewed
16	161	Bwera Katojo	Shifted
17	783	Nyamambuka	Died
18	474	Nyamambuka	Died
19	539	Nyamambuka	Shifted
20	780	Nyamambuka	Shifted
21	469	Nyamambuka	Operates a business across the border in Congo and does not back in the evening.
21	476	Karambi	Outright refusal
22	483	Karambi	Died
23	793	Karambi	Shifted
24	477	Karambi	Shifted
25	492	Karambi	Operates a business across the border in Congo and resides there
26	494	Karambi	Shifted
27	481	Karambi	Shifted
28	484	Karambi	Died
29	397	Rukoma	Mentally unstable
30	433	Rukoma	Migrated to Bunyoro
31	452	Rukoma	Shifted
32	376	Rukoma	Shifted
33	348	Rukoma	Migrated
34	465	Rukoma	Migrated
35	438	Rukoma	Migrated
36	442	Rukoma	Migrated
37	805	Kweyamba	Respondent is hospitalized.
38	804	Kweyamba	Travelled and no one knows when he is likely to come back
39	794	Kweyamba	Outright refusal
40	796	Kweyamba	Travelled, no replacement in the HH
41	21	Kweyamba	Respondent died
42	103	Bwera Katojo	Shifted to kiryandogo
43	119	Kabirizi	Shifted to Karuma
44	191	Kabirizi	Respondent died and no one to participate in the survey
45	197	Kabirizi	Shifted to Gulu
46	581	Rukoma	Shifted to Westnile
47	587	Rukoma	Shifted to Busoga
48	666	Rukoma	Respondent died, and the husband shifted
49	677	Rukoma	Shifted to Kampala
50	715	Rukoma	Shifted to Congo
51	10	Kweyamba	Outright refusal
52	200	Kabirizi	Outright refusal

4.2.7. Field challenges - Ghana

In view of the fact that all necessary arrangements and preparations were made before the commencement of fieldwork, as well as the engagement of only experienced staff members on the project, no serious problem was encountered on questionnaire administration. However, we still encountered some problems, these were some problems encountered:

- In General farmers were not at home, they were on their farms harvesting, hence we have to wait and go to the field or homes in the evening, this impacted on time.

- GPS capturing; most of the interviewers couldn't capture the GPS after they finished the interviews and because of the GPS capturing issue, some interviews timing was wrong, and others too was reading 12 hours as time spent for the interview. The most striking GPS issues was 15 cases that couldn't sync because the GPS was not capturing.

- Some common issues like death of a respondent, a respondent has moved from the location and a respondent had travelled and he is not coming back again were issues experienced across locations.

- There was a death in Wayub, but we were able to interview the wife who is now farming on the land in N-nalog in his place. We also had three deaths, two in Nayili village and one in N-nalog (the village Chief), their HH_IDs are 318, 328 and 302 respectively. One respondent also travelled in Gbadagbam, with two respondents moved from the same Gbadagbam and Jakpom, with their respective HH_IDs as follows: 411, 385 and 245.

- In Yendi, the control cluster had a lot of issues, out of the 387 respondents we were able to achieve only 371 respondents with 16 of them had various issues. The issues striking in Tusani where we have as many as 11 households have travelled and new members of their households have to be interviewed in their place. In another form Fusheina Wahabu, with HH_ID 580 was in Kpatia in 2015 but now in Tusani and was interviewed in Tusani.

- We had two deaths, 11 people moved out the village and 3 travelled and no refusal in the Yendi cluster. The Two deaths were Fuseina Mohammed and Arishetu Zakaria with HH_IDs, 668 and 619 and they were in Gundogu and Kulpanga respectively.

- The GPS readings nearly marred the effective field data collection we had, where the data guys couldn't locate 15 of the respondents we claimed to have synced. It is good we were able to resync them on time for the data team to start cleaning.

All these contributed to some delays we had on the field. By and large, we would like to say categorically that this Survey of Farming Households, 2017 – Phase two in has been professionally executed in line with the project instructions, specifications and MRS code of conduct in Ghana.

4.2.7.1. Attrition cases – Ghana

These are the respondents that had some issues and we couldn't interviewed them.

CLUSTER	HH_CODE	ITEM	DEAD	MOVED	TRAVELLED	REFUSED	TOTAL
SABOBA	385	GBADAGBAM		1			1
	411	GBADAGBAM			1		1
	245	JAKPOM		1			1
	318	NAYILI	1				1
	328	NAYILI	1				1
	302	N-NALOG	1				1
SALAGA	37	KPEMBE				1	1
	175	KPEMBE				1	1
	5	KPEMBE			1		1
	42	KPOLO		1			1
YENDI	460	BINI		1			1
	648	GUNDOGU		1			1
	436	GUNDOGU		1			1
	668	GUNDOGU	1				1
	681	KULPANGA		1			1
	631	KULPANGA			1		1
	615	KULPANGA			1		1
	627	KULPANGA		1			1
	680	KULPANGA			1		1
	619	KULPANGA	1				1
	504	KUSHEGU		1			1
	509	KUSHEGU		1			1
	510	KUSHEGU		1			1
	736	NALONGU		1			1
	778	PION		1			1
	492	ZAKOLI		1			1
			5	**14**	**5**	**2**	**26**

4.2.8. Field challenges - Benin

We faced some unprecedented challenges with the fieldwork partner in Benin. Dalberg Research often works with numerous service providers for a wide range of surveys. It is worth noting in Benin, we had worked with the same partner in 2015 and therefore we were confident that the partner would deliver on the project specifications. However, for various unexplained reasons, there were miscommunications with the cooperative presidents, delays in data collection but the most outstanding being that we were unable to account for some surveys in Benin.

Fieldwork partner unaccountability: Even though this is not usual and also given the fact that this was the same partner we worked with in 2015. We were disappointed that 7 cases in Benin were not accounted for and unfortunately our fieldwork partner was unable to provide reason for these cases being unaccounted for.

Other challenges included the following;

Duplicate IDs: In Benin, we faced cases of duplicate IDs, where a wife and husband would be interviewed. In this case, we took the interview of the person who was interviewed in 2015. The teams ended up with duplicate IDs because of any of the fowling reasons.

- Double issuance of the HH contact information sheet such that two enumerators ended up with the same HH contact Sheet
- Enumerator error in keying In the Unique ID.

Respondents not willing (10 surveys started but respondents refused along the way): This we considered to be a very high number, especially from farmer respondents who had been willing to take part in the survey and gave their consent. This was mainly from the control area in Benin. Some excused themselves and indicated that they had other commitments and would return to be interviewed but they never did. Unfortunately, we never really got to the bottom of this issue and therefore we were unable to get the reasons for the respondents stopping interviews half way.

Farmers who were unavailable/could not be traced/not picking calls were 34 cases.

4.2.8.1. Attrition cases – Benin

	Unique HH ID	Cluster/Co-op	Reason
1	15	Calavi coop	Does not pick up his phone, not found at home, could not be traced
2	107	Calavi coop	Does not pick up his phone, not found at home, could not be traced
3	115	Calavi coop	Unavailable, travelled
4	118	Calavi coop	Unavailable, not at home
5	129	Calavi coop	Unavailable, not at home
6	131	Calavi coop	Unavailable, not at home
7	134	Calavi coop	Unavailable, not at home
8	140	Calavi coop	Unavailable, not at home
9	212	COMADOV-GP coop	Very aggressive, ask never to contact him again
10	259	COMADOV-GP coop	Refuses to collaborate
11	417	APlahoué:	Unavailable, not at home
12	438	Aplahoué	Does not pick up his phone, not found at home, could not be traced
13	471	Aplahoué	Does not pick up his phone, not found at home, could not be traced
14	478	Aplahoué	Unavailable
15	502	Aplahoué	Does not pick up his phone, not found at home, could not be traced
16	514	Aplahoué	Does not pick up his phone, not found at home, could not be traced
17	582	Aplahoué	Does not pick up his phone, not found at home, could not be traced
18	596	Aplahoué	Does not pick up his phone, not found at home, could not be traced
19	597	Aplahoué	No longer farming, become a forest officer
20	638	Aplahoué	Does not pick up his phone, not found at home, could not be traced
21	646	Aplahoué	Does not pick up his phone, not found at home, could not be traced
22	654	Aplahoué	Does not pick up his phone, not found at home, could not be traced
23	667	Aplahoué	Unavailable
24	670	Aplahoué	Does not pick up his phone
25	698	Aplahoué	Unavailable
26	707	Aplahoué	Unavailable
27	719	Aplahoué	Does not pick up his phone, not found at home, could not be traced
28	470	Aplahoué	Does not pick up his phone, not found at home, could not be traced
29	602	Aplahoué	Unavailable, not at home, not picking his call
30	649	Aplahoué	Unavailable, not at home, not picking his call
31	806	Aplahoué	Unavailable, not at home, not picking his call
32	807	Aplahoué	Unavailable, not at home, not picking his call
33	810	Aplahoué	Unavailable, not at home, not picking his call
34	811	Aplahoué	Unavailable, not at home, not picking his call

4.2.9. Field challenges - Benin

We did not face major challenges except for internet access challenges which led to delays in synching completed surveys.

GPS Capture; we also had areas within the survey target areas where the team members were unable to capture the GPS readings. We later opened up the location variable in the script and had the readings keyed in manually to facilitate synching of the surveys.

5. Conclusion

Although the fieldwork in the five countries Kenya, Uganda, Mali, Benin and Ghana was faced by a number of challenges, we were able to effectively implement the data collection activities albeit not within the allocated time in Benin. We visited all the targeted respondents, and by the end of the fieldwork we had achieved a total of 3,840 out of the targeted 4002 interviews.

It is our sincere hope that the quality of the work delivered by Dalberg Research in this assignment will be up to the standard expected by 2SCALE IFDC and that the final clean data set sent to AIR for reporting will be found to be valid, accurate, and reliable and an actual reflection of the prevailing situation on the ground within the cooperatives. We will welcome any follow up queries and clarifications on the survey process and the captured primary data. For every such case, we will ensure that we provide adequate and timely responses.

We would like to acknowledge the diligent and valuable input we got from the 2SCALE Team and the AIR Team during the implementation of all the survey activities. We are particularly very grateful for the timely cooperation and support we received from among others, Rai Nisha and Juan Bonilla especially during the script testing, training and set ups, Ruth Kamunya and Ernest Archeampong's presence during the training and pilot de-brief in Kenya, Uganda and Ghana was very invaluable in providing survey-specific perspectives on a number of issues raised by the training participants, especially with reference to the survey questionnaire. It is our hope that we will have more opportunities in the future to share our skills, experiences and expertise with this and other similar dedicated team(s) from IFDC at large and 2SCALE in specific.

We would also like to register our appreciation to all the participants who contributed in one way or the other to the success of the Endline survey. We are grateful to the data collection team made up of the supervisors, team leaders and enumerators who were very dedicated in their work, the several challenges highlighted above notwithstanding. We are kindly indebted to all Dalberg Research project staff from Field, Data Processing, and Client Service Departments for their unity of purpose in the study, and to the entire Dalberg Research Management, ably led by Dr. Jasper Grosskurth and Jasper Gosselt for ensuring that there was timely and adequate institutional support during the implementation of the survey.

AIR®

AMERICAN INSTITUTES FOR RESEARCH®

Established in 1946, American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research on important social issues and delivers technical assistance, both domestically and internationally, in the areas of education, health, and workforce productivity.

MAKING RESEARCH **RELEVANT**

AMERICAN INSTITUTES FOR RESEARCH
1000 Thomas Jefferson Street NW
Washington, DC 20007-3835 | 202.403.5000
www.air.org

LOCATIONS

Domestic: Washington, DC (HQ) | Monterey, Sacramento, and San Mateo, CA | Atlanta, GA | Honolulu, HI | Chicago and Naperville, IL Indianapolis, IN | Metairie, LA | Waltham, MA | Frederick and Rockville, MD | Chapel Hill, NC | New York, NY | Columbus, OH | Cayce, SC Austin, TX | Reston, VA

International: El Salvador | Ethiopia | Haiti | Honduras | Kyrgyzstan | Tajikistan | Zambia

3852_03/18