Evaluation of the impacts of a soil fertility training project on farm productivity in the Volta region of Ghana

Kwaw Andam, International Food Policy Research Institute, Accra, Ghana

Simrin Makhija, International Food Policy Research Institute, Washington, DC, USA

David Spielman, International Food Policy Research Institute, Washington, DC, USA

Grantee Final Report

Accepted by 3ie: January 2019

International Initiative for Impact Evaluation

Note to readers

This final impact evaluation grantee report has been submitted in partial fulfilment of the requirements of grant TW4.1022 awarded under Thematic Window 4. 3ie is making it available to the public in this final report version as it was received. During the course of this evaluation, the implementing agency changed their targeting strategy without informing the research team. This created a number of challenges for the analysis, including reducing the effective sample size. As a result, one cannot be very confident that the study was able to accurately measure the effects of the intervention.

The 3ie technical quality assurance team comprises Diana Lopez-Avila, Mark Engelbert, Stuti Tripathi, Deeksha Ahuja, Samidha Malhotra, Emmanuel Jimenez, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Marie Gaarder.

3ie received funding for the Agricultural Innovation Evidence Programme from Alliance for Green Revolution in Africa, the Bill & Melinda Gates Foundation, the International Fund for Agricultural Development and the UK Department for International Development. A complete listing of all of 3ie's donors is available on the 3ie website.

Suggested citation: Andam, K, Makhija, S and Spielman, D, 2018. *Evaluation of the impacts of a soil fertility training project on farm productivity in the Volta Region of Ghana*, *3ie Grantee Final Report.* New Delhi: International Initiative for Impact Evaluation (3ie).

Acknowledgements

The authors would like to thank Africare and its staff for collaborating with us on this evaluation. Their cooperation and support throughout the study were invaluable. We would also like to thank the Alliance for a Green Revolution in Africa (AGRA) for continuous support and feedback on our work, the International Initiative for Impact Evaluation (3ie) for funding and for technical review and support throughout the study, and the CGIAR Research Program on Policies, Institutions, and Markets (PIM) for additional financial support.

This study would not have been possible without the administrative support of IFPRI's Ghana Strategy Support Program (GSSP) in Accra. We are also grateful to the Institute of Statistical Social and Economic Research (ISSER), our data collection partner and Festus Kwabena Annor-Frempong, Department of Agricultural Economics and Extension at the University of Cape Coast for taking the lead on qualitative data collection and analysis. We also extend our thanks to Benjamin Afful Jr. from University of Cape Coast and Hlupeki Phiri, for their research and logistical support. All errors are our own.

Executive summary

In 2015, Africare began its second phase of an ambitious project to promote the adoption of a suite of integrated soil fertility management (ISFM) technologies and practices to smallholder farmers in the Volta Region of Ghana. Phase 1, conducted in 2013-15, laid the foundation for the project by setting up three one-stop centres for agricultural inputs and outputs and training farmers on ISFM practices in three districts. Phase 2 expanded the training project to three additional districts. The project's main modality was to provide training to farmers via agricultural extension agents through a training-of-trainers (ToT) intervention, with the ultimate goal of increasing yields of three major crops and increasing farm incomes for 20,000 members of farmer-based organizations (FBOs) within three years.

IFPRI worked with Africare to evaluate the impacts of this intervention on a wide range of outcomes including awareness and knowledge of these ISFM practices, as well as changes in yield and income. A difference-in-difference approach with matching techniques was used as the primary evaluation method, augmented by in-depth qualitative analysis.

Baseline data were collected in early 2016 and endline data in late 2017 to evaluate the impacts of Africare's ISFM intervention over a two-year period. Africare's intervention focused on FBO members and was implemented in a manner that aimed to generate extensive spill overs to non-participating households residing in close proximity to these FBO members. Data collection focused on randomly selected households that were FBO members in the six districts where Africare was operating, as well as similar households in six comparable and adjacent control districts that were identified at project inception based on similarities in agroecological and demographic characteristics. This was supplemented with continuous collection and analysis of Africare's own project monitoring data and a qualitative study based on key informant interviews and focus group discussions conducted at midterm.

Findings from all three data sources and conversations with Africare suggest that the intervention did not reach its intended beneficiaries in the manner originally proposed. Specifically, Africare was unable to reach its intended number of beneficiaries by exclusively targeting members of FBOs in the project area. As a result, Africare made mid-course corrections in its targeting strategy and reached out extensively to non-members.

This shift in targeting had profound effects on the design, implementation, and analysis underlying IFPRI's evaluation. Importantly, the change in the population of interest rendered the sample of households that were randomly selected for the evaluation's surveys at baseline and endline unrepresentative of the population actually reached by the intervention. The change in the population of interest also resulted in much lower exposure to Africare's intervention within that same sample of households. Despite these challenges, our report attempts to analyse the extent of Africare's reach within our sample and estimate its impacts. We focus our analysis on outcome variables early on in the theory of change (what we describe as "first-order outcome variables"). Given the mid-course targeting shift, as might be expected, we are unable to find any positive impacts of the intervention on first-order outcome variables such as awareness, knowledge, and adoption of ISFM. We also do not find impacts on second-order outcome variables that are further down the theory of change, such as labour use and crop yields.

These results draw attention to critical importance of coordination between intervention design and evaluation design. Specifically, the key lesson that has emerged from this project is the mutual need to design a project with evaluation in mind while also designing an evaluation with a project in mind. In this particular case – and despite best efforts made on the part of all organizations involved – efforts to design the project and its evaluation in a carefully coordinated manner fell short. As a result, little can be said about the impact of Africare's project. It is possible that the project had an impact on its targeted households that simply could not be captured by the measurements used here. It is also possible that the project had little or no impact on such households, as the analysis in this report suggests.

Contents

List of figures and tables

Abbreviations and acronyms

AEA	agricultural extension agent
AGRA	Alliance for Green Revolution in Africa
DID	difference-in-difference
FBO	farmer-based organization
GDP	gross domestic product
GHC	Ghanaian cedi
GLSS	Ghana Living Standards Survey
ISFM	integrated soil fertility management
IFPRI	International Food Policy Research Institute
OSC	One Stop Centres
ISSER	Institute for Statistical, Social and Economic Research
MoFA	Ministry of Food and Agriculture
SLM	sustainable land management
SSA	Africa south of the Sahara
T&V	training and visit
ToT	training of trainers

1. Introduction

In Ghana, as across much of Africa south of the Sahara (SSA),[1] low agricultural output and yields are a primary reason why households remain in poverty (Molini and Paci 2015). Low agricultural output and yields affect not only on-farm productivity but are also an indirect constraint on the growth of agro-industries and the rural non-farm economy, thereby hindering both agricultural and economic transformations (Diao et al. 2018). Low agricultural productivity is partly a symptom of scarcities in the natural resources required for farming, including scarcities in soil nutrients. Thus, efforts to improve soil fertility are key to improving rural livelihoods and food security.

Agricultural scientists and development practitioners generally agree that there are readily available technical fixes for low soil fertility that are appropriate to the agro-ecological and socioeconomic context of rural SSA. One such fix is integrated soil fertility management (ISFM), which is defined as a flexible set of economically and socially acceptable uses of existing resources in conjunction with organic and mineral inputs to increase productivity (Vanaluawe 2004). Prior agronomic research on ISFM has demonstrated its efficacy under both research-station and farmer-managed trials (Vanluawe et al. 2005; Place et al. 2003).

Although information on ISFM is readily available to agricultural experts, and although its positive benefits are well established in the scientific literature, the dissemination of information on ISFM practices to farmers has been either less than desired or less successful than desired in SSA. This suggests that there is much room to improve efforts designed to increase the adoption of ISFM practices in SSA. However, there remain questions about how best to go about disseminating information on ISFM.

Conventionally, dissemination of ISFM information is considered the role of public extension agents in many SSA countries, especially given that the non-rival, non-excludable nature of the information rarely attracts private sector providers. But there is much evidence demonstrating that public extension is increasingly constrained in its ability to deliver the requisite advisory services to poor and highly fragmented smallholders (Haug 1999; Anderson 2007; Davis 2008; Birner et al. 2009). Traditional extension services have largely focused on the provision of basic inputs such as improved cultivars, inorganic fertilizer, and agrochemicals, but are less equipped to provide information on relatively complex practices such as ISFM.

There is a growing body of evidence that examines the impact of alternative approaches to promoting such complex technologies, including evaluations of various extension approaches such as farmer field schools (Davis et al. 2012; Bonan and Pagani 2017), enhanced training and visit (T&V) systems (Kondylis et al. 2014), and other demand-driven extension services (Klerkx and Leeuwis 2008), or through behavioral and social channels such as individual learning dynamics and peer effects (Hanna et al. 2014; Conely and Udry 2001). However, conclusions from these studies are almost always specific to context—to the farmers, crops, agro-ecological conditions, infrastructure, institutions, markets, and policies that influence the target population.

Consider two recent examples to illustrate this point. Kondylis et al. (2014) found that variations on the T&V extension approach in Mozambique yielded positive outcomes in terms of imparting

[1] In this report, we use the abbreviation “SSA” to denote sub-Saharan Africa. Note, however, that the Africa Union’s recommended terminology is “Africa south of the Sahara.” In documents made available in the public domain, IFPRI complies with the Africa Union’s recommendations as a matter of organizational policy.

new knowledge about sustainable land management (SLM) practices to farmers, but were less effective in increasing SLM adoption rates. Similarly, Bonan and Pagani (2017) found that the farmer field schools for children (Junior Farmer Field Schools) extension approach was effective in raising the children's and their household's knowledge of beneficial agricultural practices, but caution that this acquired knowledge did not necessarily translate into adoption.

This study seeks to add to the evidence on the impact of extension approaches used to promote the uptake of complex natural resource management practices among small-scale, resource-poor farmers in developing countries. It does so by evaluating the impact of a training of trainers (ToT) approach to promoting ISFM on input use, management practices, productivity, and household welfare in the Volta Region of Ghana.

Seven primary outcome variables were laid out in the study's pre-analysis plan: awareness and understanding of ISFM, adoption of ISFM, labour use, yields, farm profits, farm assets, and food security. The primary research questions were generated to measure the impact of Africare's intervention on these seven outcome variables. The data from the household survey and from Africare's monitoring reports reveal that only a small fraction of our sample received the treatment. This has serious implications for our impact analysis. Given this finding, we limited our analysis to four outcome variables – awareness, adoption, labour, and yield; and excluded higher order outcome variables – profits, assets, and food security.

This report is organized as follows: Section 2 describes the project intervention, the theory of change, and the research hypotheses. Section 3 lays out the context within which Africare implemented the intervention and the rationale for study site selection. Section 4 delineates the timeline of the intervention and the evaluation. Section 5 describes the evaluation design and identification strategy. Section 6 expands on the design, methods, and implementation of the project. Section 7 delves into treatment uptake and examines preliminary impacts on short-term outcome variables. Section 8 discusses the challenges encountered and preliminary findings and Section 9 makes project recommendations.

2. Intervention, theory of change and research hypotheses

2.1 The intervention

In broad terms, Ghana's agricultural sector has been characterized for several years by low productivity, limited input use, low soil fertility, land degradation from shifting cultivation ("slash-and-burn") practices, and low yields for most food staple crops. To address these issues, Africare's intervention in the Volta Region of Ghana promoted the adoption of a suite of ISFM technologies and practices that ultimately aimed to increase on-farm yields and the farm incomes of smallholders. Africare's project specifically promoted the sustainable intensification of maize, cowpea, and cassava cultivation with the provision of information on (1) production inputs and their use, (2) integrated soil management practices, and (3) marketing strategies and services. This information intervention was provided by agricultural extension agents (AEAs) employed by the Ministry of Food and Agriculture (MoFA) and trained by Africare and was targeted to members of farmer-based organizations (FBOs) in selected districts of the Volta Region.

The project was the second phase of a long-term intervention that started in 2011 under the Soil Health Programme of the Alliance for a Green Revolution in Africa (AGRA). The first phase

(2011–2014) promoted integrated soil fertility management (ISFM) in the Hohoe, Jasikan, and Kadjebi Districts in the Volta Region. The overall goal of the first phase was "to increase agricultural productivity in sustainable ways through increased use of improved ISFM technologies and significant reduction in the current practice of slash and burn agriculture in the Volta Region, a key breadbasket of Ghana" (AGRA 2014). The target for the project was to reach about 30,000 smallholder farmers by the end of the project in 2014. According to AGRA's evaluation at the end of the first phase, it was difficult to ascertain whether this goal was achieved, although it is estimated that more than 18,000 farmers adopted improved practices under the project, leading to yield increases in some cases that more than double the pre-project yield (AGRA 2014). Absent a rigorous evaluation, however, the project's achievements in the first phase could not be independently verified.

The key components of Africare's project evolved based on the needs and priorities of district agricultural extension offices and FBOs. This section describes the intervention since inception of the second phase of the project in September 2015. A more detailed discussion of the intervention follows in Section 6.

Africare's project revolved around the implementation of a series of ToT sessions for AEAs, followed by support to AEA's farmer outreach activities. These outreach activities targeted farmers to provide a combination of (1) lectures on ISFM and related topics and (2) demonstration plots and related interactions. The mode of delivery was interactive and participatory rather than instructional so as to maintain the attention and interest of farmer-participants.

The primary training topics identified and prepared were based on content from scientific documents and training guides provided by the Council for Scientific and Industrial Research (CSIR), Soil Research Institute (SRI), the Alliance for a Green Revolution in Africa (AGRA), the Africa Soil Health Consortium (ASHC), and other sources (see, for example, this handbook by Fairhurst (2012) from the ASHC).

2.1.1. Lectures on ISFM and related topics

Integrated soil fertility management: This training introduced AEAs (and, subsequently, farmers) to ISFM principles and practices. The training was designed to first gauge farmers' level of understanding of ISFM and their current practices. It then introduced ISFM as a term used for different methods and practices that are collectively designed to enhance soil fertility. The training then explored the causes of soil fertility loss and methods to detect fertility losses, with focus placed on recognizing soil erosion, nutrient deficiencies in plants, and yield trends from a given plot over time.

The training covered a range of agricultural practices which, when implemented together, could improve soil fertility. These practices included: use of green manure or farmyard manure; use of chemical/synthetic fertilizer in prescribed quantities; composting; growing cover crops; growing legumes; crop rotations with maize, cassava, and legumes; retaining and mulching of crop residues; moving away from shifting agriculture practices; encouraging the presence of specific insects, worms, and other organisms in the soil to help with decomposition and drainage; careful weed management; use of improved seeds; use of herbicides and weedicides; maize line planting; contour ploughing; and timely and accurate implementation of these recommended practices.

Soil testing: This training introduced AEAs (and, subsequently, farmers) to methods used to test nutrient and physical deficiencies in soils. Lessons included: basic principles of soil health; simple field-testing techniques such as visual inspection, colour comparisons, and tactile analysis; more complex laboratory-testing techniques; interpretation of results from these testing techniques; and appropriate responses to deficiencies in soil fertility.

Marketing, product quality, and standards: As a follow-on to the training sessions on ISFM and soil testing, Africare conducted ToTs on crop marketing, product quality, and standards. The goal of these trainings was to help improve farmers' ability to sell their produce in local markets, to aggregators, or to other market agents in a manner that would secure higher per-unit values than most farmers had been receiving.

2.1.2. Demonstration plots and related interactions

Demonstration plots: To augment the trainings described above, Africare assisted FBOs in establishing a demonstration plot within close proximity of their members' homes. Demonstration plots were donated by FBO members or other community stakeholders, and FBO members participated in the preparation, planting, management, and harvesting of crops cultivated on the plot. As is standard practice with most demonstration plots, each plot was designed to provide side-by-side comparisons of crops cultivated with and without ISFM treatments. Participating farmers were encouraged to observe and analyse differences in plant health, crop establishment, yield, and other variables. This experiential learning approach was designed to encourage individual and collective ownership of information and knowledge generated by the learning experience, and to further encourage participants to experiment with ISFM on their own farms.

Farmer field days: Africare organized farmer field days to bring together FBO members from several communities. The aim of these farmer field days was to encourage farmers to observe and participate in the actual implementation of recommended practices on a demonstration plot, and to share information on their own farming practices with a group of peers larger than their immediate FBO.

Continuous interactions with farmers: Africare and MoFA encouraged farmers to bring their farming problems and challenges to the attention of AEAs. Farmers interacted with AEAs during their visits to FBOs and communities, or they visited AEAs in their local offices or contacted them via phone to obtain feedback on implementation of recommended practices.

The Volta Region experiences two rainy seasons. A longer one approximately from April – July (major season) and a shorter one from September to November (minor season). Phase II of Africare's intervention was launched in September 2015, during the minor season. Phase II was implemented over three years and completed by mid-2018. The baseline survey collected recall data on major and minor season 2015 and the endline focused on minor season 2016 and major season 2017.

2.2 Outcomes and impacts of interest

The evaluation sought to measure the extent to which smallholder farmers in the Volta Region of Ghana benefited from Africare's training sessions on ISFM. The primary evaluation question was whether Africare's training interventions on ISFM resulted in increases in (1) awareness, learning, uptake, and adoption of selected ISFM-related technologies, (2) land and labour productivity, (3) returns to and incomes from farming, and (4) household food security and welfare.

The main outcomes and impacts of interest were reflected in the original primary research questions as follows:

1. Did Africare's ISFM training increase the awareness and use of purchased inputs and ISFM management practices and marketing strategies among smallholders?
2. Did Africare's ISFM training result in increases in land productivity (yields) for major crops and increases in labour productivity on the farm and in market participation?
3. Did Africare's ISFM training result in an increase in the returns to farming and improvements in household welfare?
4. To what extent did Africare's ISFM training result in heterogeneous outcomes (measured in terms of changes in awareness, adoption, productivity, market participation, income or welfare) among specific groups within Africare's target population?
5. Was the ISFM training designed by Africare a cost-effective means of achieving the project's desired outcomes?

The primary research questions listed above defined the variables that the study set out to measure. Section 7 explains in detail how the intervention's outreach deviated from reaching the originally defined target population, and the consequences of this deviation for the evaluation. Specifically, because the majority of farm households in the original sample never received the treatment because of the deviation, we are required to focus on the first four outcome variables and exclude higher-order outcome variables.

Outcome variables for the hypotheses listed above are constructed as follows:

1. Awareness and understanding of ISFM: The level of awareness about the different aspects of ISFM are measured using variables k1a–k1i in the questionnaire provided along with this report. In addition, we modify variables k2a–k2i to determine whether the respondents had heard about the different components of ISFM prior to Africare's intervention. Relatedly, the level of understanding of ISFM concepts is measured using a content knowledge test developed by the evaluation team in consultation with Africare. Tests are scored against 25 questions where respondents received 1 point for every correct answer and no points for a wrong answer or a "don't know" response. The 25 questions that comprise the knowledge test correspond to variables s1–s18, s20–s24, s25 (s25a, s25b, s25c), and s26 collected as part of the household questionnaire. In addition to analysing the overall score, we also analyse each individual component of the score.

2. Adoption of ISFM: ISFM is a suite of technologies which, when adopted together, have been shown to have beneficial effects on farm productivity and soil fertility. Respondents were asked to report on the use of these practices at the plot level. We have created multiple indicator variables that take on the value 1 if the respondent has adopted a given practice on any of their maize plots in the major season, and 0 otherwise.

3. Labour use: We measure labour use in a sex-disaggregated manner as man-days divided by crop area and woman-days divided by crop area for maize in the major season.

4. Yields: We rely on farmer-reported harvest quantity and area cultivated for maize in the major season to calculate yield.

2.3 Theory of change

Africare's theory of change is depicted in the following results-based framework.

Table 1: Theory of change

<table>
<tr><th colspan="3">Goal: To contribute to increased food security and farm incomes in the Volta Region of Ghana</th></tr>
<tr><td>Objective 1:
To improve input and output market services to farmers through the scale out of One – Stop – Centres (OSCs) across 5 Districts in Volta Region</td><td>Objective 2:
To improve stakeholders' capacity to sustainably deliver Integrated Soil Fertility Management (ISFM) technologies to smallholder farmers</td><td>Objective 3:
To improve the documentation & dissemination of project results &best practices</td></tr>
<tr><td>Outcome:
Increased capacity of smallholder farmers to coordinate and engage input & output markets</td><td>Outcome:
Increased adoption of ISFM technologies by farmers</td><td>Outcome:
Increased access to project information on results and best practices</td></tr>
<tr><td>Outputs:
• 2 additional OSCs established in target districts
• Established OSCs managed by private sector
• Input and output market services delivered to farmers through OSCs</td><td>Outputs:
• ISFM technologies delivery partners identified and trained
• Basic tools and equipment for ISFM technologies scaled out provided to stakeholders
• Farmers educated and trained in the use of ISFM technologies and soil testing.
• All relevant ISFM technologies transfer stakeholders actively participate in project implementation</td><td>Outputs:
• Platforms for information and knowledge sharing developed
• Functional database of farmers developed
• Database of ISFM technologies transfer stakeholders established
• Database of ISFM technologies effective in the Volta Region established
• Project impact evaluation conducted</td></tr>
<tr><td>Activities:
• Establish One Stop Centres (OSCs)
• Identify & recruit private sector to manage OSCs
• Sensitize farmers to increase access to available agricultural services
• Upgrade existing OSCs to include mechanization services
• Facilitate the development of a Terms-Of-Reference for the management of OSCs
• Link farmers to Micro Financial Institutions</td><td>Activities:
• Train FBOs on ISFM technologies
• Institute a project steering committee
• Establish ISFM demonstration plots
• Number of Extension Agents trained in ISFM technologies to provide support to farmers
• Organize farmer field days
• Organize farmer exchange visits
• Conduct soil analysis
• Provide soil testing kits</td><td>Activities:
• Conduct baseline on target communities
• Develop M&E tracking systems/tools for information gathering
• Organize stakeholder fora to share information
• Organize regular radio programmes to share information & educate the pubic
• Create platforms for information & knowledge sharing, including ICT.</td></tr>
</table>

• Train farmers on market requirements • Train farmers in produce quality & standards		• Create a project database • Conduct third party project impact evaluation

The primary theory of change underlying the intervention was that improved access to information and best practices for smallholder farmers would lead to increased adoption of ISFM technologies and practices. The training was therefore aimed at building capacities of farmers to coordinate and engage in input and output markets and the OSCs eventually serve as focal points for efficient post-harvest handling and marketing. The newly adopted ISFM technologies and practices would contribute to increased agricultural productivity and efficient agricultural hubs, which in turn will cause improved food security and income, thereby reducing poverty.

A secondary theory of change assumes that evidence garnered from the intervention can be used to advocate for improvements in the enabling environment—public policies to promote ISFM through innovative extension activities, investment incentives for private service providers to supply ISFM products/services, or resources for training of service providers—that would further accelerate the capacity of public, private, and civil society organizations to replicate and scale-up similar interventions.

This evaluation focused on project objective 2: improve stakeholders' capacity to sustainably deliver ISFM technologies to smallholder farmers. The activities listed under objective 2 include training FBO members and AEAs on ISFM. Through the ToT project, Africare sought to train AEAs who would in turn train FBO members on ISFM. Implicit in this logic is the assumption that information on ISFM will pass from Africare to the FBO members *via AEAs* with minimal dilution. Findings from the qualitative study suggest that AEAs were indeed trained on ISFM practices and AEAs reported being familiar with the ISFM content taught to them. The study also found that the objectives of the ToT were achieved for the most part – AEAs believed that they had successfully transferred knowledge on the technologies and practices to farmers. However, the ToT has failed to create awareness on soil testing, marketing, and produce quality and standards.

Africare's activities listed under objective 2 directly translate into several of its expected outputs, including (1) training partners identified and trained, (2) basic tools and equipment scaled out, and (3) farmers trained in soil testing and ISFM.

The theory of change posits that the intervention's outputs (especially the output on "farmers trained in soil testing and ISFM") will lead to the desired outcomes (especially "increased adoption of ISFM technologies by farmers") if and only if farmers are aware of and understand ISFM and its component parts. This theory of change is a fairly standard reflection of accumulated experience with, and evidence on, learning and innovation processes in agriculture, and is described in such seminal works as Griliches (1957), Rogers (1962), and Schultz (1975), among many others (see reviews by Feder, Just, and Zilberman (1985) and, more recently, Jack (2013)). In its simplest sense, these studies define adoption as the sustained use of a given technology, with "final adoption at the level of the individual farmer… defined as the degree of use of a new technology in long-run equilibrium when the farmer has full information about the technology and its potential" (Feder, Just, and Zilberman 1985). In the context of this evaluation, however, we define adoption as an initial trialling of the technology given the relatively short timeframe of the intervention itself.

Despite the role that new technologies and technical change can play in increasing agricultural productivity and farmer welfare, technology adoption has been fraught with challenges. There is a rich literature to illustrate the many barriers to adoption (Foster and Rosenzweig 2010; Feder, Just, and Zilberman 1985; Jack 2013). Identified barriers include insufficient access to information and lack of knowledge (Foster and Rosenzweig 1995; Munshi 2004), liquidity constraints and lack of trust (Cole et al., 2013), wealth and credit constraints (Giné et al. 2008; Hill et al. 2013), risk (Feder 1980; Dercon and Christiaensen 2011), and farm size and land tenure (Feder, Just, and Zilberman 1985). These challenges to adoption pose a potential threat to Africare's theory of change and the transformation of outputs into outcomes.

3. Context

Ghana is an ideal country for testing the effectiveness of a public intervention to improve agricultural productivity and reduce rural poverty, especially an intervention focused on ISFM. The country has already made significant strides in reducing rural poverty. According to data from the Ghana Living Standards Surveys (GLSS), the rural poverty rate was nearly halved between 1991/92 and 2012/13, falling from around 65 percent to 38 percent. Much of this reduction was, however, achieved as a result of the movement of rural labour out of agriculture and into the services sector in Ghana's urban areas. Thus, while Ghana attained lower-middle-income status in 2010 and experienced rapid economic growth on the order of 10 percent per year in 2011–2013, agricultural growth has been much slower at 2.9 –percent during the same period. By 2014–2016, the economy's growth rate had dropped to 3.8 percent per year while the agriculture growth rate moved up to 3.6 percent. The drop in the gross domestic product (GDP) growth rate exposed the country's vulnerability to global commodity price volatility, while the increase in the agricultural sector's growth rate masks considerable year-on-year volatility. This suggests that underinvestment in agricultural productivity growth – including growth in regions dominated by lower-value staple crop production – remains a major challenge to Ghana's national strategy for economic growth and development.

To be sure, many of Ghana's policymakers recognize the challenges associated with advancing agricultural productivity growth. The sector employs around 40 percent of the labour force in Ghana and contributes 20 percent to the country's GDP. However, farms have seen negligible increases in agricultural productivity over the past two decades. For example, yields for maize, a crop grown in all parts of the country and one of the focal crops for this project, stood at about 1.7 metric tonnes per hectare and have not increased appreciably in spite of several efforts to introduce improved varieties and other modern inputs throughout the country.

Ghana's investments in agriculture have also not included a meaningful focus on natural resource management, including ISFM, despite the potential economic benefits such practices provide to smallholder farmers and, in the longer term, to the environmental sustainability of intensive production systems. The main policy action in the past decade to address soil fertility was a fertilizer subsidy programme, which started in 2007 and by 2010 represented around a quarter of the budget of the Ministry of Agriculture (MoFA). The only other notable government programmes for improving soil fertility have been sporadic public campaigns against bush-burning, shifting agriculture, and slash-and-burn practices. Throughout much of Ghana, farmers still tend to increase on-farm production by expanding land under cultivation (rather than intensifying production on existing land) by clearing virgin lands. Basic management principles related to enhancing soil fertility, such as crop rotations and crop residue retention, may be known by many farmers, but they are rarely put into practice. This continues despite the fact that land is becoming increasingly scarce, especially in the southern part of Ghana. This includes the Volta Region, where Africare's project was implemented.

3.1 Participant recruitment

Participants in Africare's ISFM project in the Volta Region were initially selected based on (1) geographic location, (2) membership in a farmer-based organization, and (3) willingness of participants to engage with the implementing organization and its partners. The *a priori* importance of criterion (3) to Africare effectively ruled out scope for randomly assigning the intervention to households and/or farmer-based organizations in the project area.

Geographic locations (districts) for project implementation were determined purposefully by Africare in consultation with government partners and other stakeholders. In Africare's words,

> "The selection of the targeted districts for scale out was carefully done based on farmer populations, poverty and unemployment rates, and readiness of stakeholders to participate in the scale out project."

'Readiness of stakeholders to participate in the scale-out project' refers to the willingness of district agricultural officers to participate in Africare's intervention; specifically, the willingness of for the district agricultural offices that are responsible for AEA activities and performance to allow their AEAs to be trained by Africare and, in turn, have those AEAs train farmers in their district on ISFM practices. This authorizing environment is a potentially important and early step in ensuring project success.

Africare's "treatment" districts were Jasikan, Kadjebi, Hohoe, Afadjato South, North Dayi, and Kpando. Farmers in these districts were randomly selected for participation in both the baseline and endline surveys.

The "control" districts were purposively selected from among similar districts in the Volta Region that (1) adjoined the treatment districts and (2) shared similar agro-ecological and socio-economic characteristics. The choice to select a control group from outside the treatment districts rather than draw a control group from untreated households within the treatment districts was made to minimize the potential for (1) contamination of the controls (control households directly receiving the treatment) and (2) spill overs affecting the controls (control households indirectly receiving the treatment from the treated). The resultant control group also allowed the evaluation to identify any systematic and observable differences between the farmers chosen by Africare and other, similar farmers.

The choice of control districts required careful consideration and was only finalized during a brainstorming session in December 2015 that was attended by district directors of agriculture, agricultural extension agents, Africare staff, and the evaluation team. The question put to participants was "which districts are most similar to the treatment districts in terms of agricultural potential, population, and demographic characteristics?" In addressing the question, the team also made use of official data on the districts in the Volta Region. The team benefited from the experience of the MoFA staff, who had historical knowledge of the agricultural trends in the districts in the years preceding the project. The control districts thus selected were Biakoye, South Dayi, Ho Municipal, Ho West, Kranchi East, and Nkwanta South. Farmers in these districts were randomly selected for participation in both the baseline and endline surveys. The strategy adopted to select control districts was the best option available given the circumstances. It was the most effective way to identify the control group given the constraints on the evaluation design imposed by the intervention's targeting strategy.

3.2 External validity and representativeness of the sample

A few words on the *external validity* of this study are in order. There are two main contributions that this impact evaluation could make to the existing body of knowledge on information dissemination, agricultural extension, and technology adoption. First, in a narrow sense, one might be interested in the implications of these findings for efforts to promote ISFM practices among smallholders in Ghana and in other, similar countries and smallholder production systems in SSA. This implies that the evaluation contributes to answering questions about whether existing or future projects that rely on a ToT approach are likely to experience similar outcomes in awareness and adoption of ISFM.

Secondly, in a broader sense, one might be interested in the implications of these findings for efforts to promote a broader class of natural resource management principles and practices among similar populations, countries, and contexts. This suggests that the evaluation contributes to answering questions about whether existing or future projects – whether using a ToT approach or some other approach – are likely to experience similar outcomes in awareness and adoption for technologies and practices such as the proper application of seeds, fertilizer, pesticides, crop management techniques, water management practices, or other inputs. In other words, is the level of success that was achieved by Africare in promoting ISFM indicative of the goals that are achievable for other, similar agricultural development projects in general?

Establishing the implications of the findings for populations, markets, and agro-ecologies throughout Ghana requires some consideration of the representativeness of the study's sampled population. The sample were drawn from FBO-member households in rural districts in the Volta Region. Given the almost-universal implementation of the intervention among farmers in FBOs in the selected project districts, and since the region is predominantly rural, it is reasonable to assume that the farmers were fairly representative of households in the Volta region. But to assess the sample's representativeness vis-à-vis the entire country, we need to compare the region's profile to the rest of the country. We do this by drawing on data from the most recent round of the Ghana Living Standards Survey (GLSS), a nationwide survey, as well as several other sources.

In Ghana, poverty is concentrated in the three northern regions (Northern, Upper East, and Upper West Regions), where the poverty incidence ranged from 44 to 71 percent in 2013, compared to a national average of 24 percent. Among the seven remaining southern regions, the Volta Region had the highest level of poverty in 2013 at 34 percent, and the data show that the poor in the Volta Region are living in deeper poverty than they were in 2006 (Cooke, Hague, and McKay 2016). Agricultural wages, however, have grown slightly faster in the Volta region than in any other region except for the Western Region, and the region has seen the highest share of government spending on agriculture, although the region's share of the national crop area is the second-lowest in the country (Diao et al. 2018).

Thus, in terms of representativeness, the Volta Region falls between the poorest northern regions and the relatively wealthier southern regions in its poverty status. But in other measures – including key measures of socioeconomic and agro-ecological characteristics – it is also somewhat distinct from many other regions of the country. Thus, from both the evaluator's and the policymaker's perspective, there is only limited external validity to the rest of Ghana from the sample drawn for this study.

4. Timeline

IFPRI's team made its first scoping visit in June 2015 during which it met with Africare's field coordinator to get a better understanding of Africare's interventions piloted in Phase 1 and its plans for Phase 2. It was Phase 2 that was to be the subject of this evaluation, and in September 2015, Africare officially launched Phase 2 of its ISFM intervention. On a second trip in December 2015, IFPRI's team briefed Africare on the impact evaluation and pre-tested the household and community questionnaires. In February–March 2016, IFPRI collected baseline data from households and communities ensuring to sample from villages that Africare had not reached yet. In October–November 2017, IFPRI collected endline data. See Table 2.

Table 2: Timeline

Year	Month	Project duration	IFPRI Prep	IFPRI Baseline	IFPRI Qual	IFPRI Endline	Seasons
	June						Major
	July						
	August						
2015	September						
	October						Minor
	November						
	December						
	January						
	February						
	March						
	April						
	May						Major
	June						
2016	July						
	August						
	September						
	October						Minor
	November						
	December						
	January						
	February						
	March						
	April						
	May						Major
	June						
2017	July						
	August						
	September						
	October						Minor
	November						
	December						

5. Evaluation: Design, methods and implementation

5.1 Ethical review

This evaluation underwent rigorous ethical review for human subject research both by IFPRI and by the University of Ghana. Specifically, the study was first reviewed by IFPRI's Internal Review Board and then reviewed by the Ethics Committee for the Humanities at the University of Ghana. The study was approved by both ethical review boards and received subsequent (continuing) approvals that covered the entire study duration. All approval documents are available on file.

To meet IFPRI and the Ethics Committee's standards and requirements for ethical human subject research, appropriate measures were taken to obtain informed consent from study participants. Personal information collected from consenting respondents during all surveys was carefully managed to ensure privacy; respondents' time was carefully respected by limiting the survey to the shortest possible duration; and a token cash compensation of GHC 12 (~USD 3) was provided to respondents for participating in the survey, which is about 50 percent greater than the national minimum daily wage.[2] Tablets assigned to survey enumerators for data collection in the field were password protected, and identifying information in the resulting datasets was separated from any and all files subsequently used for data analysis. Only selected researchers on IFPRI's team have access to data files containing personal information.

5.2 Evaluation design/identification strategy

A randomized controlled trial would likely have provided the most rigorous basis on which to explore a causal relationship between the treatment and the outcomes associated with Africare's intervention. However, as noted earlier, Africare and its partners selected the project sites based on observable factors. Specifically, participants in Africare's ISFM project in the Volta Region were initially selected based on (1) geographic location, (2) membership in a farmer-based organization, and (3) willingness of participants to engage with the implementing organization and its partners.

A difference-in-difference (DID) approach with matching techniques was chosen and set forth in the pre-analysis plan. The DID approach is based on the principle that a simple before-after comparison of outcomes for the trained farmers may be biased because of unobserved factors that affect outcomes and that changed along with the project. However, if these unobserved factors also affected the control farmers, then DID can remove the bias and isolate the treatment effect. In addition, by using data from both treatment and control farmers before and after the intervention, the DID approach seeks to remove any bias from permanent differences between the two groups.

Given that Africare's a priori selection of project districts and participant self-selection as part of its implementation strategy effectively ruled out a randomized controlled trial, the DID approach and matching techniques provide the next-best means of creating a credible counterfactual with which to compare treatment effects on "treated" participants.

Thus, the study initially set out to make comparisons between two groups:

1. Treatment group: FBO members who reside in one of the six districts where Africare-trained AEAs conducted training sessions
2. Control group: FBO members who reside in one of six districts where Africare-trained AEAs did *not* conduct training sessions

As set forth in the pre-analysis plan, a DID approach with matching techniques was to be used to identify and quantify the causal relationship between participation in ISFM training sessions and outcome variables related to increases in (1) awareness, learning, uptake, and adoption of

[2] In January 2016, immediately prior to baseline data collection, the Government of Ghana announced a National Daily Minimum Wage of GHC 8. Thus, the compensation provided for the survey was 50 percent above minimum wage.

selected ISFM-related technologies, (2) land and labour productivity, (3) returns to and incomes from farming, and (4) household food security and welfare. This combined approach allows for comparison of treated households against a set of similar but untreated households that are identified based on observable characteristics, with comparisons being made both before and after the intervention. This approach offers a straightforward means of evaluating the impact of Africare's project when compared against similar households that were not part of the project (with vs. without), and by controlling for exogenous changes affecting both project participants and non-participants (before vs. after).

As a first step in this study's analysis of the intervention's impact, we explore results based only on the DID approach. This is done because of concerns raised about changes to the treatment assignment during the project's implementation that may threaten the ability to identify an area of common support through matching. The causes and consequences of changes in the treatment assignment are explained in greater detail below.

5.3 Sample size calculations

Despite Africare's project documents indicating high expected rates of adoption, yield increases, and income gains from ISFM, more conservative estimates of outcomes were used to calculate the sample size requirements for this evaluation. The power calculations were based on data for the Volta Region from the Ghana Living Standards Survey (GLSS) Round 5 conducted in 2005–2006 and Round 6 conducted in 2012–2013. The GLSS is a nationally representative household survey with a comprehensive agriculture module. These data were a second-best option to agricultural census data or sample survey data from the project area itself. Despite this, they do provide information on production, income, and fertilizer use, which are, in turn, used in the power calculations.

Given the assumptions used in the power calculations (see Online appendix D for details), combined with the underlying distributional properties of the data, conservative expectations of project impact, and evaluation budget constraints, the evaluation required 665 farmers in the treatment arm (T) with 67 clusters (villages) and 10 farmers per cluster.

For the control arm, given that this evaluation relied on propensity score matching, a larger number of FBOs, farms, and farmers was required to optimize the match between treatment and control. Estimates of the proportion of farmers that fell into the region of common support were used to determine how much larger the control group must be. This estimate was calculated as 0.97, after which the size of the control group for a pure RCT (665) was divided by this proportion to arrive at a control group size of 686 households with 69 clusters and 10 farmers per cluster. Thus, the total sample size determined was 1,360 households.

As per standard practice, the power calculations were conducted before baseline to determine the sample size needed to measure minimum detectable effects. Given that Africare's shift in targeting came to light well after baseline, no modifications could be made to the power calculations.

Table 3: Sample size calculations

Arm	Clusters	Farmers per cluster	Total
Treatment	67	10	670
Control	69	10	690
Total	**136**	**-**	**1360**

5.4 Sampling Frame

5.4.1 Quantitative

The sampling frame was designed as follows. First, it was determined that the population of interest was to be FBO members, since these individuals were the primary beneficiaries of extension services provided by AEAs trained by Africare using the ToT approach. To minimize the possibility of within-community spill overs from such services, the sampling cluster was chosen at the community level.

Second, Africare and MoFA officials provided a complete listing of all FBOs and their membership size in the districts where Africare's project was operating. These districts are Jasikan, Kadjebi, Hohoe, Afadjato South, North Dayi, and Kpando. From this list, all FBOs that participated in Africare's earlier (Phase 1) activities and its most recent (Phase 2) activities were dropped from the sampling frame due to the possibility of pre-project exposure to the treatment.

Third, a total of 115 communities were initially identified in these treatment districts. The list was further pared down based on the previously conducted power calculations. Specifically, communities with fewer than 10 members in an FBO were dropped, resulting in 112 communities from which to sample. Probability proportional to size sampling was then used to draw a sample of 63 communities.

Fourth, six control districts were purposefully chosen to determine whether they were proximate to and possessed observable and generally similar agro-ecological and socioeconomic characteristics to, the districts where Africare's project was operating. These control districts were chosen based on multiple consultations conducted with district directors of agriculture, their AEAs, and Africare staff based in the project region. The control districts are Biakoye, South Dayi, Ho Municipal, Ho West, Kranchi East, and Nkwanta South.

Fifth, each of the (control) districts was visited and a listing exercise was conducted for all FBOs and data on their membership were collected. A total of 78 communities were initially identified in these control districts. The list was further pared down (again, based on prior power calculations), resulting in 75 communities. Given that 69 communities were to be sampled from 75, to overcompensate for the need for a significantly large set of controls for propensity score matching, all 75 communities were selected.

At the time of data collection, enumerator teams were asked to verify the FBO member listing in each community and then draw a random sample of 15 FBO members. They were to survey the first 10. If a potential respondent did not cultivate maize, cassava, or cowpea, the enumerators were to use one of the five replacement households.

During survey implementation, enumeration teams were able to conduct surveys in 59 treatment communities and 75 control communities. In total, data were collected from 1,333 households at baseline, 629 in the treatment group and 704 in the control group. At endline, follow-up data were collected on 1,292 households, 608 in the treatment group and 684 in the control group. The survey attrition rate is extremely low at 3 percent.

Of the three focus crops, maize is the primary crop cultivated by our sample. The results from the qualitative study also find that majority of smallholders in this region grow only maize. Thus, we restrict our analysis to farmers who cultivated maize.

5.4.2 Qualitative

The main respondents for the qualitative study were (1) Africare staff who facilitated the implementation of ISFM, (2) AEAs who had been trained under the ToT approach by Africare, (3) AEAs who had not been trained by Africare, (4) directors of the Departments of Agriculture in the districts where the projects were carried out, and (5) smallholder farmers, including FBO leaders, who were the beneficiaries of the ISFM project. Different procedures were used to select key informants from each respondent category, as explained below.

Selection of Africare staff: The four Africare project staff (the two technical staff members in charge of the two zones – Kpando municipal, Afadjato South district, and North Dayi district zone and Hohoe municipal, Jasikan district, and Kadjebi district zone), the chief executive, and the technical staff member in charge of business development) were purposively selected as respondents for the qualitative study. They were used as key informants to confirm the project objectives, Africare's ISFM technologies dissemination, the extension tools used, and as a link to the extension agents and farmer groups.

Selection of MoFA extension agents (AEAs): The extension agents were selected in a slightly different manner. A census of the agricultural extension agents (those trained by Africare and others) and the district directors at the Departments of Agriculture involved in the project at Jasikan, Kajebi, Kpando, Hohoe, Afadjato South, and North Dayi districts were engaged through focus group discussions to examine their perception of the ISFM technologies, the extension methodologies, and challenges with the implementation of the project. A total of six focus group discussions involving 43 extension agents and directors of the Departments of Agriculture were held during the study. Three extension agents were purposively selected for the personal interviews based on the FBOs interviewed in the districts/municipalities. For every FBO selected, the AEA in charge of the area was selected. This was corroborated with the training and methodologies AEAs indicated to have used. Eighteen AEAs were interviewed separately.

Selection of smallholder farmers: The smallholder farmer respondents were selected purposively to ensure that the effectiveness of extension methodologies (project delivery) was assessed from the perspectives of farmers who had participated in the establishment of "One Stop Centres" (OSCs), demonstration plots, farmer field days, collaboration with esoko (a mobile service provider providing support to rural producers and consumers), and radio programmes on Sekpele 104.3 FM. The selection of smallholder farmers therefore started with the purposive selection of communities where various extension methodologies were used in each district or municipality. A community was then randomly selected out of the list of communities collected from Africare where each methodology had been used. Different focus group discussions were therefore conducted for different farmer groups based on the extension methodology.

Three individual farmers (including one executive of the FBO, a female, and a male) were selected for in-depth interviews to elicit information on project rationale, effectiveness of delivery, ownership, and challenges with adoption of ISFM technologies after each focus group discussion. Two other communities from the adjoining districts, namely Aboabo in the Biakoye district and Tsokpokope in Afadjato South district, were purposively selected to examine if awareness and knowledge on ISFM has diffused into other districts. Focus group discussions were held with the farmers at Aboabo and Tsokpokope and followed with the personal interviews. Overall, 19 focus group discussions were conducted involving 289 smallholder farmers during the study, and 63 individual farmer interviews were conducted.

5.5 Data

Baseline data were collected from an individual survey of 1,333 households and a community survey of 134 communities from March to April 2016. This was prior to the intervention being launched in most treatment communities, and prior to farmers' application of any information obtained from the treatment in their fields. The baseline survey was designed by IFPRI in late 2015 and early 2016. Pre-testing was conducted in February 2016 and the survey itself was conducted in February-March 2016. The survey comprised a household-level and a community-level instrument. The former was administered to household heads/primary decision makers in 1,333 households in the sampled districts based on the sampling frame described above. The latter was administered to groups of no more than 10 knowledgeable community members in 134 communities.

The individual questionnaire contained the following modules and relied on recall for questions on input use and harvested quantities for the Major and Minor seasons in 2015.

1. Primary male and female decision maker details
2. Maize, cassava, and cowpea plot details
3. Crop inputs by season
4. Crop production by season
5. Other plots – size and harvest
6. Cost of crop inputs by season
7. Labour use by season
8. Seed input by season
9. Crop sales by season
10. Crop storage by season
11. Interaction with agricultural extension agents and Africare
12. Non-farming income sources
13. Credit access and savings
14. Assets – farm and non-farm
15. Food security
16. Food consumption
17. Non-food consumption expenditure
18. Recent shocks and household welfare
19. Content knowledge test on ISFM
20. Risk preference game

The community questionnaire contained the following modules:

1. Roster of informants
2. Access to basic services
3. Agricultural labour, extension services, and agricultural problems faced by community
4. Land use
5. Demographics and land details
6. Water access, shocks, and food consumption
7. Market prices and conversion of non-standard commonly used units
8. Experience with Africare

Endline data were collected following the major agricultural season in October–November 2017. Data were collected on the same modules as the baseline survey and relied on farmer recall for input use and harvest quantity for minor season 2016/17 and major season 2017. At endline, a brief survey of extension agents was also carried out to collect data on their workload and salaries. The survey instruments can be found in Appendix A.

The Institute for Statistical, Social and Economic Research (ISSER) at the University of Ghana served as the data collection partner for both baseline and endline surveys for the entire sample of respondents in the treatment and control districts. IFPRI and ISSER together trained enumerators hired by ISSER on the survey instrument and how to administer it in the local language (Ewe). At both baseline and endline, the instruments were pre-tested in Ewe-speaking communities outside the study area. Data were collected using SurveyCTO on Android tablets.

Respondents were compensated GHC 12 (~USD 3).

Given that the survey was administered using SurveyCTO, multiple checks were in place to ensure that responses to certain questions were within a reasonable range, question flow was logical, and all questions were mandatory. ISSER used its team of supervisors, project coordinators and managers to keep a close check on the data. Enumerators and supervisors checked the data before sending it to the server and project coordinators took trips to the field to audit their work.

5.5.1 Avoiding bias

Multiple measures were put in place to minimize and avoid bias in the data collection process. For example, enumerators were trained in interviewing techniques to avoid bias associated with interviewer effects. Consideration was also given to contamination and spill over effects. To prevent contamination, farmers in the control group were chosen from districts that were not covered by the Africare project. However, spill overs from treated farmers to other, untreated farmers in the project area is an outcome of interest but unfortunately could not be measured due to resource constraints.

In addition to bias from spill overs, it was important that the evaluation be free from any significant biasing effects from participants altering their behaviour due to the study, i.e., reactivity effects. In this evaluation, we do not consider John Henry or Hawthorne effects a major cause for concern. For more details on issues related to avoiding bias, please refer to Appendix I.

5.5.2 Data quality control

One of the primary advantages of using SurveyCTO in this study was the minimization of errors that otherwise occur when data from paper surveys are entered to a digital interface. To further ensure collection of high quality data, a two-way data feedback system was put in place between ISSER's enumeration teams in the field and IFPRI's research team in Accra and Washington, DC. Every week the ISSER team would send a report of the households from which data were collected and IFPRI would share a list of households whose data had reached the SurveyCTO server. These two lists were then corroborated to ensure that data were being collected and submitted in a timely manner. In addition, IFPRI created a data dashboard to monitor key variables being collected by ISSER's enumeration teams. For example, careful monitoring was conducted for unique id entry, data from the risk preference game, units of quantity measurement, and other key variables in the survey to ensure that responses were reasonable. In cases of unreasonable responses, duplicate household ids, and other errors, a report was sent immediately to ISSER. Appropriate corrections were then agreed upon between ISSER and IFPRI, and IFPRI entered the corrections into the dataset manually using Stata.

Appropriate measures were also put in place to uniformly clean and code the data. Price information from the community survey was used to calculate values for consumption and fertilizer. Using community level price information helps reduce the possibility of outliers in individual responses. All binary variables that were coded with ones and twos in the questionnaire were recoded to ones and zeros for "yes" and "no" responses respectively. To ensure certain plot-level variables were comparable across respondents, the quantities were consistently divided by total crop area. Missing values were not imputed.

6. Project design, methods and implementation

The second phase of the project, which is the subject of this evaluation, started in 2015. By the end of the project, Africare estimated that the following outcomes would be achieved:

- A 70 percent increase in the number of farmers recording increased food security and incomes
- Yield increases on the order of 213 percent for maize, 188 percent for cassava, and 400 percent for cowpea
- 17,000 farmers with access to production information and best practices
- 16,000 farmers educated and trained in the use of ISFM technologies
- 15,000 farmers with access to and participation in input and output markets
- 15,000 farmers adopting ISFM technologies
- 6,000 hectares of farmland under ISFM

To achieve these goals, Africare implemented an intervention that consisted of two components. The first component aimed to improve marketing services through privately-managed OSC agro-input retailing and FBO-managed warehousing services. The second component aimed to provide training on ISFM through training of trainers at an OSC training centre and other convenient locations. This evaluation focused on the second component,[3] and therefore the remainder of this section delves further into the ToT.

The content of the ToT on ISFM included the following: (1) causes of soil fertility loss and methods to detect fertility losses such as soil erosion, nutrient deficiencies in plants, and yield trends from a given plot over time; (2) practices that can improve soil fertility, namely use of green manure, farmyard manure, and chemical/synthetic fertilizer, composting, growing cover crops, growing legumes, crop rotations with maize and legumes, retaining and mulching of crop residues, moving away from shifting agriculture practices, encouraging the presence of specific insects, worms, and other organisms in the soil to help with decomposition and drainage, careful weed management, use of improved seeds, use of herbicides and weedicides, maize row planting, contour ploughing; and (3) timely accurate implementation of these recommended practices. Other ISFM-related practices covered in the ToT included allowing fallow periods, breaking of hard pans, avoiding slash-and-burn agriculture, soil acidity correction, harvesting of rainwater, and planting across the slope.

[3] Given the complexity of OSC establishment, operations, and management during the first phase, the evaluation team considered that it would be more realistic to focus on the second component of the project.

In addition to ISFM, the ToT project involved lectures on soil testing, marketing strategies and services, and produce quality and standards. The marketing strategies and services training focused on sub-topics such as the need for farmers to identify buyers and plan the sale of their produce before production, producing to meet the quality requirements of the customers, and marketing in groups to aid bulk buying, linkage, and attraction from aggregators for good pricing. The produce quality and standards training is aimed at producing the best quality produce to meet the standards of the customers. Sub-topics included adopting good agronomic practices such as seed selection, cultural practices, and timely harvesting; and post-harvest practices such as cleaning, drying, sorting, and packaging.

The soil testing training introduced AEAs to methods of testing nutrient and physical deficiencies in soils; principles of soil health; simple field-testing techniques such as visual inspection, colour comparisons, tactile analysis, interpretation of results from laboratory testing techniques; and appropriate responses to deficiencies in soil fertility. Over 40 AEAs and directors participated in the training sessions.

The AEAs in turn transferred the knowledge and skills on ISFM, soil testing, marketing strategies and services, and produce quality and standards obtained at the ToT sessions to FBO members and, especially in the later stages of the project, interested farmers who were non-FBO members. Continuous interactions between farmers and AEAs occurred in the farming communities, offices of the district or Municipal Department of Agriculture, and over the phone. The purpose of these interactions included arranging for one-on-one meetings to identify and solve problems and to follow up on activities. AEAs used the lecture method to convey the messages on causes of soil fertility loss, methods to detect fertility losses, marketing strategies and services, and produce quality and standards.

Africare maintained a monitoring system that included information on the number of AEAs trained, the number of farmers trained by those AEAs, the number of villages and FBOs covered by the project, the number of non-FBO farmers trained, and the number of demonstration plots. A summary of the data from this monitoring system are presented in Online appendix E. The IFPRI evaluation team used updates from the Africare monitoring system to track progress and to detect any deviations from the original implementation plan.

In addition to the regular monitoring system, the evaluation included a detailed qualitative assessment at the midterm stage of implementation, from February to April 2017. The qualitative assessment was designed to go beyond the raw data collected in the monitoring system to explore relevant underlying facets of project implementation that may not be readily observable from the monitoring data. It complemented the evaluation data, which were quantitative data collected primarily from farmers, by conducting qualitative interviews with farmers, AEAs and district directors, and the project implementers (Africare staff). Details of the qualitative sampling strategy are provided in Section 5 above. The objectives of the qualitative assessment were as follows:

1) To confirm the stated project rationale from the perspective of smallholder farmers, understand the associated expectations of opportunities, costs, benefits, and risks, and the underlying theory of change.
2) To comment on the effectiveness of project delivery to date as perceived by different stakeholders.

3) To examine the challenges faced and extent to which the project addressed them.
4) To determine the perceptions around ownership of the project, and
5) To make suggestions to improve project implementation.

6.1 Extent to which actual group of beneficiaries matched the treatment group

The project had planned to target 20,000 farmers over a three-year period (2015–2018). The project's goal was to reach out to roughly 7,000 farmers per year in six target districts of the Volta Region – Jasikan, Kajebi, Kpando, Hohoe, Afajato South and North Dayi. The project planned to train 30 AEAs in a year, each of whom are capable of providing training to between 90 and 150 farmers per semester. The project targets are shown in the table below.

Table 4: Africare project targets

District	Targets					
	3-year target	1st-year target	1st-semester target	No. of agricultural extension agents (AEAs) for 1st year	Minimum target for each AEA per semester	Maximum target for each AEA per semester
		(farmers)		(AEAs)	(farmers)	
Afadzato South	3,000	1,000	500	5	100	150
Kpando	5,000	1,700	850	6	150	200
North Dayi	5,000	1,700	850	6	150	200
Hohoe	3,000	1,000	500	5	100	150
Jasikan	2,000	700	350	4	90	140
Kadjebi	2,000	700	350	4	90	140
Total	**20,000**	**6,800**	**3,400**	**30**	**680**	**980**

Source: Africare (2016).

The following table presents data from Africare's monitoring report on farmers trained between June 2015 and December 2017.

Table 5: Aggregates of farmers reached from June 2015 to December 2017

District	Produce quality and standards training	Marketing training	ISFM training
Afadzato South	397	367	523
Kpando	1,722	1,832	2,261
North Dayi	825	909	1,096
Hohoe	1,484	1,431	3,084
Jasikan	2,517	2,648	2,808
Kadjebi	1,028	1,106	2,128
Non-FBO farmers	5,796	5,796	7,324
Total	**13,769**	**14,089**	**19,224**

Source: Collated by authors using Africare's monitoring data.

Comparing Africare's project targets with their monitoring data, we find that while their over-arching target of training 20,000 was almost met, the number of FBO members to be trained fell short. Almost 40 percent of farmers trained were non-FBO members. This poses a problem because Africare's intervention was designed to target FBO members. While the ISFM training sessions were not meant to exclude anyone (farmers who wished to sit in on a session were welcomed), Africare planned to monitor progress of the intervention for FBO members exclusively. Around two years into the project's launch of phase II, it was brought to the attention of the research team that FBO structures[4] were more or less non-existent and Africare accordingly changed their approach to targeting all farmers instead of just FBO members. Given that the study's sampling and data collection were done based on the original rollout plan that targeted FBOs, the study will underestimate the outreach and impacts of this project because it does not have data on the new target group of non-FBO farmers.

6.2 Deviation of intervention from planned activities

There were two major changes in project implementation that we consider to be potential sources of confounders in the evaluation. First, it appears the project targets in terms of number of FBOs and communities to be reached were overly ambitious. Therefore, several of the farmers in the treatment groups may not have actually participated in the project's outreach activities at any time, or they may have received the intervention a few months before the end of the project. In early 2017, barely 18 months before the expected end of the project, a review of Africare's monitoring data showed that 77 percent of communities and 88 percent of FBOs targeted by the intervention were yet to be included in the project's activities. A second major change was the inclusion of non-FBO members in the project. This seems to have been a remedial measure taken by Africare to cover more participants before the end of the project. Beyond the shift in targeting of beneficiaries, we know of no other unexpected or adverse events in the treatment group.

7. Impact analysis and results of the key evaluation questions

7.1 Quantitative specification

We employ a DID model to measure impacts on the treatment group from baseline to endline with the following equation

$$y_{i,endline} - y_{i,baseline} = \beta_0 + \beta_1 T + \beta_2 X_{\mathrm{i}} + \varepsilon_{it}$$

where *i* indexes the household and *t* indexes time (Imbens and Wooldridge 2007). *y* denotes the outcome variable of interest. The variable *T* is a binary variable that takes on the value 1 if household is treated and 0 if control, while *X* is a vector of controls. The variable ε is an iid error term.

For outcome variables where baseline data are not available, we estimate a single difference model as follows:

$$y_{it} = \alpha_0 + \alpha_1 T + \alpha_2 X_{\mathrm{i}} + \vartheta_{it}$$

In both these models, we cluster standard errors at the community level.

As described in further detail below, we further refine the comparisons of treated and control variables by using a propensity score, the probability of a household belonging to the treated

[4] FBO structures refers to the system of organizing farmers into FBOs.

group. The following equation is used to estimate the propensity score using a Probit regression:

$$PS_X = \Pr(T = 1|x) = Probit(Pi) = \gamma_0 + \gamma_1 Z + \varepsilon_{it}$$

In the specification above, PS_X stands for the propensity score, which equals the conditional probability, *PR(T=1|X)*, that a household will be assigned to the treatment group, operationalized by the Probit regression in which *Z* is a vector of observed covariates. We used the following observed covariates in the propensity score estimations: household head characteristics (age, received education, received vocational training, sex), nonfarm asset index, nonfarm income, household applied for a loan, household is Ewe speaking, household has a savings account, household size, and distance to market. The variable ε is an iid error term.

7.2 Balance test

As noted earlier, Africare's intervention was carried out in six districts of the Volta Region, and six comparable districts were identified to serve as the control. Using this district-level distinction between treatment and control, Table 6 below presents a difference-of-means test.

Table 6a: Balance test

Variable	Total	Control	Treatment	Diff
Baseline: Age of household head	48.95	48.3	49.61	1.304
	[12.4]	[0.73]	[0.70]	[1.004]
Baseline: Head received any education	0.72	0.68	0.76	0.080**
	[0.45]	[0.03]	[0.03]	[0.039]
Baseline: Head received vocational education	0.02	0.02	0.03	0.005
	[0.15]	[0.01]	[0.01]	[0.009]
Baseline: Ewe is main language	0.74	0.7	0.78	0.077
	[0.44]	[0.05]	[0.04]	[0.064]
Baseline: Household size	6.04	6.45	5.62	-0.829***
	[3.25]	[0.19]	[0.13]	[0.234]
Baseline: Fraction of dependent members	0.40	0.41	0.38	-0.024
	[0.23]	[0.01]	[0.01]	[0.016]
Baseline: Male head of household	0.88	0.89	0.87	-0.021
	[0.32]	[0.01]	[0.02]	[0.025]
Baseline: Area(acres) maize (.=0)	6.67	7.11	6.23	-0.878
	[32.74]	[0.88]	[1.98]	[2.157]
Baseline: Distance to market (minutes)	24.96	26.25	23.66	-2.593
	[23.36]	[1.68]	[1.92]	[2.540]
Baseline: Applied for a loan in the last 12 months	0.32	0.29	0.36	0.067**
	[0.47]	[0.02]	[0.03]	[0.033]
Baseline: Has savings with a bank	0.41	0.42	0.4	-0.023
	[0.49]	[0.03]	[0.02]	[0.034]
Baseline: Nonfarm asset index	0.03	0.09	-0.03	-0.118
	[0.99]	[0.07]	[0.05]	[0.089]
Baseline: Total non-farm income	2917.90	3270.33	2563.06	-707.274
	[9796.52]	[588.35]	[336.30]	[675.656]
Baseline: Total consumption expenditure GHC (annual)	31216.29	29127.84	33319.06	4191.224
	[31072.35]	[1300.15]	[2256.35]	[2593.953]
Baseline: Score on knowledge test (on 25)	10.83	10.39	11.27	0.876***
	[3.10]	[0.19]	[0.19]	[0.265]
N	1170	587	583	1170

Significance levels: * < 10% ** < 5% *** < 1%

In column (1) standard deviations in parenthesis. In columns (2)-(4) robust standard errors in parentheses, clustered at the community level. The sample used for these balance tests includes maize-growing farmers for whom there were no missing observations.

Table 6b: Balance test

Variable	Total	Control	Treatment	Diff
ISFM practices				
Baseline: Ploughed land	0.80	0.86	0.73	-0.133***
	[0.40]	[0.02]	[0.03]	[0.030]
Baseline: Applied manure	0.33	0.31	0.36	0.045
	[0.47]	[0.02]	[0.03]	[0.033]
Baseline: Applied other organic fertilizer	0.35	0.31	0.38	0.074**
	[0.48]	[0.02]	[0.02]	[0.033]
Baseline: Grows leguminous trees	0.28	0.27	0.28	0.005
	[0.45]	[0.02]	[0.02]	[0.029]
Baseline: Grows leguminous plants (cowpea)	0.08	0.1	0.05	-0.042**
	[0.27]	[0.02]	[0.01]	[0.019]
Baseline: Grows leguminous plants (groundnut)	0.07	0.09	0.05	-0.030*
	[0.26]	[0.01]	[0.01]	[0.018]
Baseline: Used purchased seed	0.28	0.29	0.26	-0.034
	[0.45]	[0.02]	[0.02]	[0.035]
Baseline: Used weedicide	0.58	0.63	0.52	-0.117***
	[0.49]	[0.03]	[0.03]	[0.043]
Baseline: Used pesticide	0.03	0.03	0.03	0.009
	[0.17]	[0.01]	[0.01]	[0.011]
Baseline: Value of NPK used (by area)	29.44	33.99	24.86	-9.129
	[84.10]	[4.84]	[3.79]	[6.130]
Baseline: Maize yield (kg/acre)	329.30	336.9	321.66	-15.243
	[527.72]	[29.42]	[19.46]	[35.166]
Baseline: Maize man-days (by area)	12.79	12.96	12.63	-0.33
	[20.12]	[1.02]	[0.92]	[1.367]
Baseline: Maize woman-days (by area)	8.27	7.68	8.87	1.194
	[35.63]	[0.79]	[2.10]	[2.237]
N	1170	587	583	1170

Significance levels: * < 10% ** < 5% *** < 1%

In column (1) standard deviations in parenthesis. In columns (2)-(4) robust standard errors in parentheses, clustered at the community level. The sample used for these balance tests includes maize-growing farmers for whom there were no missing observations.

As shown in the tables above, a large number of the variables of interest are not balanced at baseline. This is not entirely surprising as Africare's assignment of the treatment was not random. As noted earlier, participants in Africare's ISFM project were initially selected based on (1) geographic location, (2) membership in a farmer-based organization, and (3) willingness of participants to engage with the implementing organization and its partners. In the following sections, we address this imbalance using two methods: redefining the treatment group and adopting matching methods.

7.3 Extent of Africare's outreach

As noted earlier, the sample of treated households for this evaluation was drawn from a list of communities that Africare intended to reach. At the outset, we analysed respondents' self-reported participation in Africare's activities to understand the reach of their activities. The table below reports the percentage of farmers, by treatment and control districts, who reported participating in Africare's activities. We find participation rates at endline to be extremely low and posit two possible causes for this.

First, based on correspondence with Africare late in the project cycle, we learned that the project's targeting strategy had shifted dramatically – from a focus on FBOs and their members, to a less targeted approach designed to recruit *any* farmers and communities that demanded their services. This is reflected in Africare's monitoring data for FBO members: only about half of the treatment communities in our sample actually received treatment. And since Africare's project monitoring system only records the number of FBO members and other farmers trained at a *community* level, it is impossible to verify if the specific farmers in our sample attended the training or not. Thus, given that the evaluation's sampling frame was designed to draw from a population of FBO members[5] – based on Africare's initial targeting strategy – we surmise that the shift in targeting strategy largely explains the very low uptake rates observed.[6]

Second, although the project rolled out in September 2015 and the baseline survey was conducted in February-March 2016, the percentage of respondents in the treatment districts reporting participation in the different components of the Africare project is relatively high. We attribute this to a challenge in framing the question to respondents correctly. Discussions with Africare indicated that many farmers were likely to be unclear about whether a project component or activity they participated in was associated with Africare. However, most would likely assume that the component or activity was sponsored by MoFA. Asking respondents if they attended/participated in a project "conducted by MoFA in association with Africare" might have yielded higher response rates if respondents conflated this project with other MoFA projects. This implies that the high uptake rates reported below are primarily attributable to measurement or reporting error. This early finding led us to revise the questions in the endline survey to ask about components of the project more generally. Unexpectedly, we find a drop in reported uptake at endline. One possible explanation for this is that at baseline respondents were referring to participation prior to September 2015. However, this seems unlikely as our sample did not include communities that were targeted by Phase 1 of Africare's project. Thus, we conclude that the responses at baseline were probably due to measurement or reporting error and we use responses at endline, instead of baseline, to measure participation.

[5] Prior to data collection we conducted a listing exercise of FBO members. We did not encounter any problems finding FBO members for the survey. A possible explanation for this is that FBOs exist on paper but the system is defunct – farmers might be FBO members but no longer organize themselves in these structures. Thus, locating individual FBO members for data collection did not pose a problem but bringing farmers together based on FBO membership was likely tricky for Africare.

[6] In fact, when we shared our endline data collection plan with Africare, we were asked to survey different communities where the trainings took place and not the communities that were surveyed at baseline as we would not find farmers there who were treated.

Table 7: Self-reported participation rates

Baseline	Treatment	Control	Endline	Treatment	Control
I have heard **radio** programmes by MoFA in association with Africare discussing issues of soil fertility on maize, cassava, or cowpea	61.68	29.77	Since September 2015, I have heard a **radio** programme discussing ISFM (integrated soil fertility management)	52.47	28.22
I have participated in the preparation of a **demonstration plot** set up by MoFA in association with Africare	26.52	6.74	Since September 2015, I have participated in the preparation of a **demonstration plot**	32.24	20.32
I attended **soil testing training** conducted by MoFA in association with Africare	19.24	2.20	Since September 2015, I attended **soil testing training**	10.86	5.99
I attended **marketing training** conducted by MoFA in association with Africare	27.35	2.79	Since September 2015, I attended **marketing of produce training**	16.78	6.43
I attended **produce quality training** conducted by MoFA in association with Africare	26.03	2.49	Since September 2015, I attended **produce quality and standards training**	12.34	3.22
I attended **ISFM training** conducted by MoFA in association with Africare	21.38	1.76	Since September 2015, I attended **ISFM (integrated soil fertility management) training**	13.65	3.22

Note: The sample used above is the entire sample of respondents for whom we have data at both baseline and endline (n=1,292).

Given these issues, it would be incorrect to define treatment status simply by whether the household resided in one of the districts where Africare implemented its project, since most of the farmers within these districts in our sample were not reached. Thus, we explore other definitions of treatment below.

1. Treatment definition 1: Monitoring match
   Using Africare's monitoring data and manually matching communities in which FBOs were trained, we find that trainings were conducted in 51 percent of the communities in our sample. Of the 608 respondents in treated districts, 308 lived in a community in which Africare trained some FBOs. This does not imply that all FBOs in the community were treated but assumes that the spill overs from the training(s) could have impacted all respondents from that community. This may be a strong assumption about spill overs extending to all respondents in a community, but our field experience in observing a training project suggests that many residents of a community do listen in on a training project taking place in the community, and that such projects are rarely conducted in a manner that excludes community members.

2. Treatment definition 2: Attended ISFM training
   The most important aspect of Africare's intervention was the training on ISFM, and the other activities and training projects conducted were all designed to support the ISFM training. In theory, this training is the primary intervention and potentially a good way to measure the extent of treatment. Unfortunately, less than 14 percent of respondents in the treatment districts reported attending a training project on ISFM.

3. Treatment definition 3: Attended any training
   In addition to the ISFM training, Africare conducted two secondary training projects, the progress of which they monitored regularly. These two training projects included: (1) produce quality and standards training and (2) marketing of produce training. Thus, another possible method of defining treatment is if a respondent reported participating in any one of these three training projects. Using this definition of treatment, 22 percent of respondents in the treated districts reported participating in at least one of the three training projects, and 9 percent of the control reported the same. The downside of this definition is that a farmer who received only marketing training cannot be assumed to know anything about ISFM through Africare's project, implying that this definition may overstate the participation rate.

4. Treatment definition 4: Self-reported participation
   At endline, respondents were also asked a series of three questions regarding Africare, namely
   i. Do you know about Africare?
   ii. Do you know about Africare's ISFM activities?
   iii. Did you participate in any of these activities?

   A mere 29 percent of respondents in the treatment districts reported participating in Africare's activities.

The table below depicts the percentage of respondents who were treated, disaggregated by treatment and control districts:

Table 8: Treatment definitions

Treated	Treated districts Control	Treatment
Definition 1: Monitoring match	0	50.66
Definition 2: Attended ISFM training	3.22	13.65
Definition 3: Attended any training	8.92	22.2
Definition 4: Self-reported participation	2.49	28.78

Note: The sample used above is the entire sample of respondents for whom we have data at both baseline and endline (n=1,292).

Given that the reach of Africare's treatment was neither universal nor consistent with the project's initial targeting strategy, using the districts to distinguish between treated and non-treated respondents may overstate the reach of the project. Given the pros and cons of these different definitions of treatment, we suggest that the best definition is Definition 1, which leverages the information generated by Africare's monitoring data and assumes spill overs within the community. While there is no credible way to test the validity of this assumption we find that this is the best option given the data we have. Additionally, we recognize that the assumption implicit in this definition of treatment is that Africare's monitoring data is reliable. We make this assumption because we have no credible reason to believe otherwise. Table 9a and 9b show results from a cross-tab between the "Control" and "Treatment" defined by the district allocation in the columns with "Control" and "Treatment" defined by the monitoring match in the rows. Table 9a and 9b show the new treatment and control group sizes using definition 1 (in bold) for the entire sample and maize sub-sample, respectively. Table 10 presents results from a (pre-matching) balance test using this definition of treatment allocation. We find better balance using this definition of treatment.

Table 9a: Treatment and control group sizes using definition 1

		Treated districts		
		Control	Treatment	Total
Treated (Definition 1: Monitoring match)	Control	684	300	**984**
	Treatment	0	308	**308**
	Total	684	608	1292

Note: The sample used for this table is the entire sample of respondents for whom we have data at both baseline and endline (n=1,292)

Table 9a: Treatment and control group sizes using definition 1 (maize sample)

		Treated districts		
		Control	Treatment	Total
Treated (Definition 1: Monitoring match)	Control	587	284	**871**
	Treatment	0	299	**299**
	Total	587	583	1170

Note: The sample used for this table includes maize-growing farmers for whom there were no missing observations.

Table 10: Balance test (treatment defined using definition 1 – monitoring match)

Variable	Total	Control	Treatment	Diff
Baseline: Age of household head	48.95	48.57	50.06	1.485
	[12.4]	[0.56]	[1.09]	[1.217]

Baseline: Head received any education	0.72	0.71	0.76	0.046
	[0.45]	[0.02]	[0.04]	[0.049]
Baseline: Head received vocational education	0.02	0.02	0.03	0.009
	[0.15]	[0.00]	[0.01]	[0.011]
Baseline: Ewe is main language	0.74	0.75	0.72	-0.028
	[0.44]	[0.04]	[0.07]	[0.078]
Baseline: Household size	6.04	6.13	5.78	-0.346
	[3.25]	[0.15]	[0.18]	[0.236]
Baseline: Fraction of dependent members	0.40	0.4	0.38	-0.025
	[0.23]	[0.01]	[0.02]	[0.019]
Baseline: Male head of household	0.88	0.88	0.88	0.001
	[0.32]	[0.01]	[0.03]	[0.034]
Baseline: Area(acres) maize (.=0)	6.67	6.19	8.06	1.867
	[32.74]	[0.62]	[3.80]	[3.794]
Baseline: Distance to market (minutes)	24.96	26.12	21.57	-4.551
	[23.36]	[1.45]	[2.54]	[2.898]
Baseline: Applied for a loan in the last 12 months	0.32	0.31	0.35	0.036
	[0.47]	[0.02]	[0.03]	[0.036]
Baseline: Has savings with a bank	0.41	0.42	0.38	-0.04
	[0.49]	[0.02]	[0.03]	[0.035]
Baseline: Nonfarm asset index	0.03	0.01	0.09	0.075
	[0.99]	[0.05]	[0.08]	[0.098]
Baseline: Total non-farm income	2917.90	3088.96	2419.6	-669.363
	[9796.52]	[419.01]	[520.55]	[662.463]
Baseline: Total consumption expenditure GHC (annual)	31216.29	31283.27	31021.14	-262.131
	[31072.35]	[1518.79]	[2687.98]	[3054.457]
Baseline: Score on knowledge test (on 25)	10.83	10.63	11.41	0.783***
	[3.10]	[0.17]	[0.22]	[0.273]
ISFM practices				
Baseline: Ploughed land	0.80	0.8	0.78	-0.02
	[0.40]	[0.02]	[0.03]	[0.037]
Baseline: Applied manure	0.33	0.33	0.34	0.014
	[0.47]	[0.02]	[0.03]	[0.039]
Baseline: Applied other organic fertilizer	0.35	0.33	0.38	0.051
	[0.48]	[0.02]	[0.03]	[0.038]
Baseline: Grows leguminous trees	0.28	0.28	0.25	-0.033
	[0.45]	[0.02]	[0.03]	[0.032]
Baseline: Grows leguminous plants (cowpea)	0.08	0.08	0.07	-0.012
	[0.27]	[0.01]	[0.02]	[0.020]
Baseline: Grows leguminous plants (groundnut)	0.07	0.07	0.07	0.005
	[0.26]	[0.01]	[0.02]	[0.019]
Baseline: Used purchased seed	0.28	0.28	0.25	-0.029
	[0.45]	[0.02]	[0.03]	[0.039]
Baseline: Used weedicide	0.58	0.59	0.52	-0.076
	[0.49]	[0.03]	[0.05]	[0.051]
Baseline: Used pesticide	0.03	0.03	0.03	-0.004
	[0.17]	[0.01]	[0.01]	[0.011]
Baseline: Value of NPK used (by area)	29.44	30.92	25.13	-5.796
	[84.10]	[3.79]	[4.83]	[6.085]

Baseline: Maize yield (kg/acre)	329.30	340.85	295.66	-45.187
	[527.72]	[22.56]	[19.41]	[29.576]
Baseline: Maize man-days (by area)	12.79	13.06	12.01	-1.056
	[20.12]	[0.82]	[1.24]	[1.466]
Baseline: Maize woman-days (by area)	8.27	8.71	7.01	-1.698
	[35.63]	[1.48]	[0.72]	[1.637]
N	1170	871	299	1170

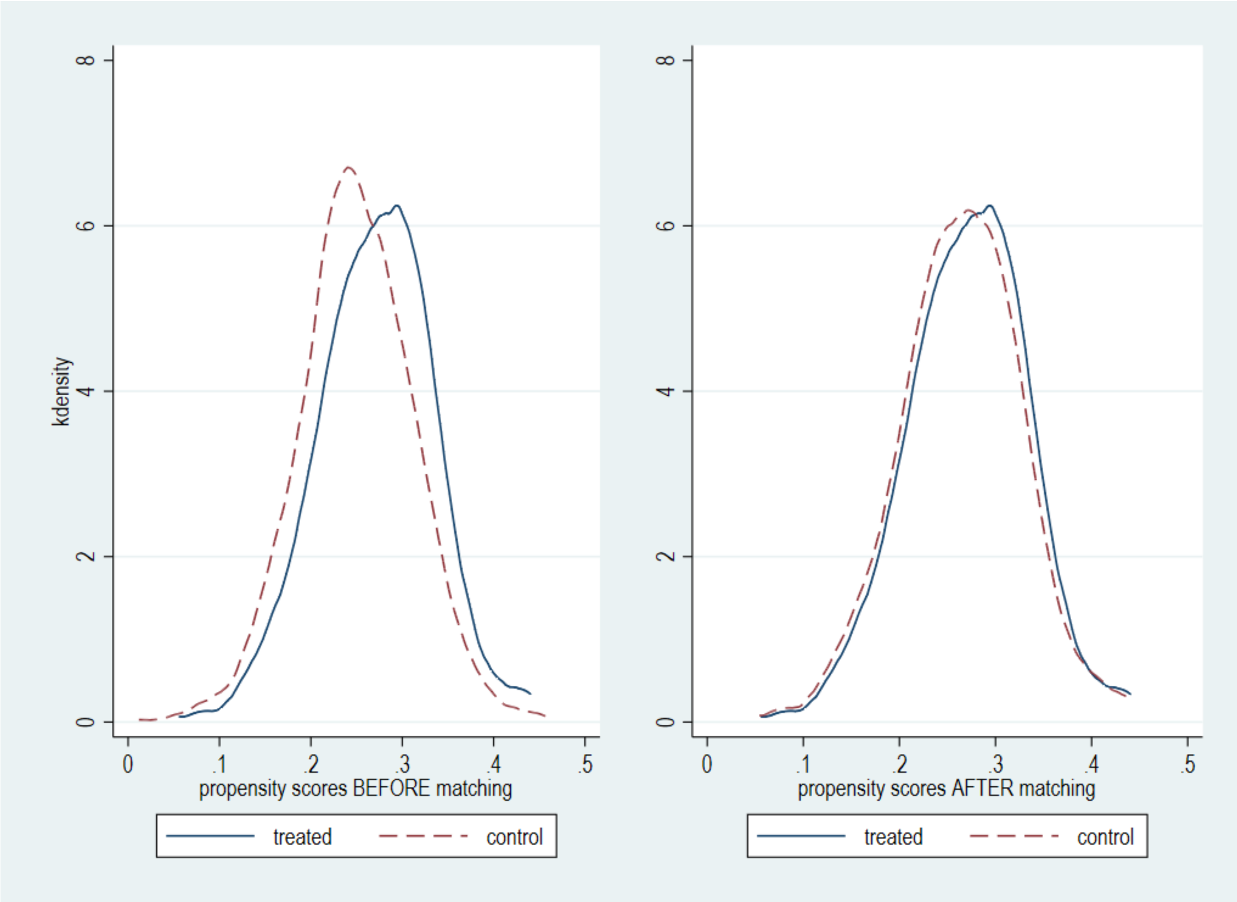
Significance levels: * < 10% ** < 5% *** < 1%
In column (1) standard deviations in parenthesis. In columns (2)-(4) robust standard errors in parentheses, clustered at the community level.
The sample used for these balance tests includes maize-growing farmers for whom there were no missing observations.

7.4 Propensity score matching

This reclassification of our treatment group in accordance with Africare's monitoring data has effectively created two control groups – a spill over control group (284) and a non-spill over or pure control group (587). We eliminate any confounding bias between these two groups using semi-parametric matching estimators based on a wide vector of pre-treatment covariates (Ho et al.; Abadie and Imbens 2016). As set out in the pre-analysis plan (see Online appendix C) we adopt propensity sore matching to establish a valid counterfactual using the households in the control (as defined by definition 1). We use a set of baseline covariates[7] that affect project participation and outcomes of project participation but only through participation in the project. As shown in the graph[8] below, we find very strong common support overlap between the treatment and control groups using nearest neighbour matching. This is further supported by the covariate balancing tests which show insignificant differences in the distribution of the treatment and control groups across pre-treatment covariates[9]. Any remaining bias that could arise from unobserved time-invariant confounders is eliminated by the difference-in-difference approach we have adopted to estimate treatment effects.

[7] The propensity score is estimated using household head characteristics (age, received education, received vocational training, sex), nonfarm asset index, nonfarm income, household applied for a loan, household is Ewe speaking, household has a savings account, household size, and distance to market.
[8] Code for this graph is from https://www.statalist.org/forums/forum/general-stata-discussion/general/1145219-psmatch2-graph-for-propensity-score-matching
[9] We use the -pstest- command on Stata to find balance in baseline level covariates after matching. Additionally, the balancing property is satisfied across 4 blocks using the -pscore- command on Stata.

Figure 1: Distribution of propensity scores before and after matching, nearest neighbour estimator



Next, we tackle the definition and measurement of key outcome variables laid out in the pre-analysis plan using this definition of treatment.

7.5 Awareness and knowledge

To measure respondents' awareness of the different ISFM practices we asked a series of questions for each practice. The ISFM practices covered included:

- Application of manure/organic fertilizer
- Application of inorganic/chemical fertilizer
- Intercropping with legumes (cowpea)
- Mulching
- Ploughing
- Use of improved seed and other planting materials
- Zero tillage
- Slash-no burn
- Good farm sanitation

For each of these practices, we asked the following two questions:

- Have you heard about the given practice?
- When did you FIRST hear about it?

For each of these practices we constructed two indicator variables to capture our "awareness" outcome. The first takes on the value 1 (0 if otherwise) if the respondent had heard about the practice, and the second takes on the value 1 (0 if otherwise) if the respondent had heard about it after minor season 2015, which was the inception season of Africare's intervention. We estimate a single difference model for these variables since the data for these questions were collected only at endline.

Table 11: Awareness about ISFM practices

	Endline value treated	Endline value control	Single difference	Nearest neighbour (k=4)	Epanech-nikov kernel	Radius
	(1)	(2)	(3)	(4)	(5)	(6)
Awareness						
Manure/organic fertilizer	0.86	0.67	0.178***	0.146***	0.0180***	0.169***
			(0.034)	(0.038)	(0.039)	(0.35)
Inorganic/chemical fertilizer	0.89	0.82	0.067**	0.058*	0.069**	0.069**
			(0.029)	(0.031)	(0.029)	(0.028)
Intercropping with legumes	0.73	0.57	0.150***	0.138***	0.149***	0.149***
			(0.035)	(0.042)	(0.032)	(0.038)
Mulching	0.69	0.55	0.136***	0.121**	0.131**	0.137***
			(0.037)	(0.058)	(0.053)	(0.035)
Ploughing	0.79	0.65	0.139***	0.108**	0.135***	0.129***
			(0.036)	(0.045)	(0.030)	(0.041)
Improved seed and other planting materials	0.81	0.63	0.167***	0.160***	0.166***	0.162***
			(0.029)	(0.037)	(0.031)	(0.025)
Zero tillage	0.37	0.29	0.072*	0.064	0.071*	0.070
			(0.042))	(0.060)	(0.043)	(0.051)
Slash-no burn	0.70	0.45	0.237***	0.236***	0.238***	0.241***
			(0.053)	(0.055)	(0.045)	(0.056)
Good farm sanitation	0.75	0.38	0.368***	0.366***	0.369***	0.365***
			(0.057)	(0.050)	(0.051)	(0.068)
Observations (treated)			299	299	299	299
Observations (control)			871	871	871	871
Observations (total)	299	871	1170	1170	1170	1170

Notes: Column (3) displays results from a single difference OLS regression with robust standard errors in parentheses, clustered at the community level. We control for the same covariates used to construct the propensity score. Estimates in columns (4) – (6) are obtained using the -psmatch2- command on Stata where standard errors are bootstrapped (50 replications) and clustered at the community level. Column (4) uses a nearest neighbour estimator with k = 4 and calliper = 0.01. Column (5) uses an Epanechnikov kernel estimator and column (6) uses a radius matching estimator with calliper = 0.01. Common support is imposed in all estimations.

Table 12: Awareness about ISFM practices after Africare's intervention

	Endline value treated	Endline value control	Single difference	Nearest neighbour (k=4)	Epanech-nikov kernel	Radius
	(1)	(2)	(3)	(4)	(5)	(6)
Awareness after project commencement						
Manure/organic fertilizer	0.07	0.08	-0.007	-0.023	-0.008	-0.008
			(0.021)	(0.025)	(0.025)	(0.024)
Inorganic/chemical fertilizer	0.05	0.06	-0.007	-0.01	-0.009	-0.009
			(0.015)	(0.021)	(0.017)	(0.020)
Intercropping with legumes	0.05	0.05	0.008	0.002	0.007	0.008
			(0.014)	(0.019)	(0.018)	(0.015)
Mulching	0.05	0.05	0.009	0.003	0.007	0.008
			(0.018)	(0.021)	(0.017)	(0.016)
Ploughing	0.03	0.01	0.018	0.017	0.017	0.018**
			(0.011)	(0.013)	(0.013)	(0.007)
Improved seed and other planting materials	0.08	0.05	0.038**	0.027	0.035*	0.031
			(0.018)	(0.023)	(0.019)	(0.021)
Zero tillage	0.03	0.02	0.012	0.013	0.011	0.01
			(0.011)	(0.010)	(0.010)	(0.010)
Slash-no burn	0.04	0.02	0.016	0.007	0.014	0.015
			(0.015)	(0.018)	(0.014)	(0.014)
Good farm sanitation	0.03	0.03	0.009	0.011	0.008	0.007
			(0.017)	(0.015)	(0.016)	(0.017)
Observations (treated)			299	299	299	299
Observations (control)			871	871	871	871
Observations (total)	299	871	1170	1170	1170	1170

Notes: This tabled adapted from the formatted used in table 7 in Bonan and Pagani (2017). Column (3) displays results from a single difference OLS regression with robust standard errors in parentheses, clustered at the community level. We control for the same covariates used to construct the propensity score. Estimates in columns (4) – (6) are obtained using the -psmatch2- command on Stata where standard errors are bootstrapped (50 replications) and clustered at the community level. Column (4) uses a nearest neighbour estimator with k = 4 and calliper = 0.01. Column (5) uses an Epanechnikov kernel estimator and column (6) uses a radius matching estimator with calliper = 0.01. Common support is imposed in all estimations.

The results in Tables 11 and 12 indicate that a larger proportion of the treatment group had heard about the different components of ISFM. However, there is no significant difference between the treatment and control groups when it comes to whether they had heard about these technologies *after* the inception of Africare's intervention. Thus, we do not have a credible basis for attributing the higher levels of awareness in the treatment group to Africare's intervention. A possible explanation for this is that Africare's radio programs were broadcasted in both treatment and control districts. This could be the driving factor behind why we do not see a significant difference between whether treated and control farmers heard about these technologies *after* Africare's intervention was launched.

Next, we explore our "knowledge" outcome variables. Knowledge about ISFM practices was measured using a 25-question content knowledge test created by the evaluation team in consultation with Africare and based on training materials provided by Africare. The test included questions on the appropriate use of chemical fertilizers, organic inputs, pests, and other inputs and management practices. A full list of questions can be found in Online appendix G.

Given that this test was administered at both baseline and endline, we estimate a DID model to measure the impacts of the project on knowledge about ISFM. We created indicator variables that take on the value 1 for a correct answer and 0 otherwise. We measure the impacts of Africare's training project on the respondents' content knowledge score (out of 25).

Using a DID estimator with and without matching we find no impacts of the ISFM intervention on treated respondents' knowledge about ISFM. Results are shown in Table 13 below.

Table 13: DID estimate of impact on understanding, adoption, labour, and yield

	Baseline value treated	Baseline value control	DID	Nearest neighbour (k=4)	Epanech-nikov kernel	Radius
	(1)	(2)	(3)	(4)	(5)	(6)
Knowledge score (max = 25)	11.41	10.63	0.1	0.001	0.101	0.073
			(0.327)	(0.395)	(0.262)	(0.331)
Adoption of ISFM practices						
Ploughing	0.78	0.80	-0.029	-0.001	-0.021	-0.019
			(0.038)	(0.048)	(0.040)	(0.036)
Manure	0.34	0.33	0.012	0.011	0.014	0.019
			(0.045)	(0.060)	(0.033)	(0.041)
Other organic fertilizers	0.38	0.33	-0.014	-0.02	-0.015	-0.007
			(0.041)	(0.052)	(0.043)	(0.043)
Leguminous trees	0.25	0.28	0.067	0.077	0.073*	0.083*
			(0.042)	(0.049)	(0.043)	(0.044)
Leguminous plants (cowpea)	0.07	0.08	-0.036	-0.02	-0.032	-0.029
			(0.028)	(0.028)	(0.032)	(0.030)
Leguminous plants (groundnut)	0.07	0.07	-0.041	-0.03	-0.038	-0.039*
			(0.028)	(0.031)	(0,025)	(0.022)
Purchased seed	0.25	0.28	0.022	0.027	0.02	0.017
			(0.036)	(0.041)	(0.032)	(0.045)
Weedicide	0.52	0.59	0.037	0.044	0.041	0.051
			(0.063)	(0.082)	(0.053)	(0.064)
Pesticide	0.03	0.03	0.069**	0.055	0.070***	0.068**
			(0.030)	(0.036)	-0.025	(0.033)
NPK Value	25.13	30.92	-9.099	-37.877	-9.467	-13.007
			(14.375)	(46.955)	(14.701)	(20.801)
Labour						
Maize man-days	12.01	13.06	0.009	-0.935	0.312	-0.215
			(2.890)	(3.185)	(2.540)	(2.488)
Maize woman-days	7.01	8.71	-0.055	-0.275	0.173	-0.272
			(2.465)	(2.286)	(2.146)	(2.265)
Maize yield	295.66	340.85	- 117.285	-190.045	-104.325	-133.397
			(91.370)	(156.694)	(78.090)	(95.751)
Observations (treated)			299	299	299	299
Observations (control)			871	871	871	871
Observations (total)	299	871	1170	1170	1170	1170

Notes: This tabled adapted from the formatted used in table 7 in Bonan and Pagani (2017). Column (3) displays results from a double difference OLS regression with robust standard errors in parentheses, clustered at the community level. We control for the same covariates used to construct the propensity score. Estimates in columns (4) – (6) are obtained using the -psmatch2- command on Stata where standard errors are bootstrapped (50 replications) and clustered at the community level. Column (4) uses a nearest neighbour estimator with k = 4 and calliper = 0.01. Column (5) uses an Epanechnikov kernel estimator and column (6) uses a radius matching estimator with calliper = 0.01. Common support is imposed in all estimations. For all estimation the dependent variable is the difference between a given outcome at endline and baseline.

Results from the qualitative study

The results from the qualitative study indicate that farmers are aware of the causes of low soil fertility and ISFM practices that were taught by AEAs through the ToT project. Quotes from FGDs and interviews show how beneficiary farmers felt about the project:

"*We applied no fertilizer, no pesticide, and we did not harvest early. We have realized after participation in the ISFM project that these were major contributors to low productivity"* (FGD with members of Winners Farmers Association in Gbefi in Kpando district).

"Even those using fertilizer did not apply the right quantities and at the right time for the crops to grow and develop well because they lacked appropriate knowledge before ISFM project". (Personal Interview with a farmer of the Dzidefo Cooperative, Marketing and Farming Society at Have Etoe in Afadjato South district).

"We slash and burn during the major season. As you are aware, this is a forested area and it is very difficult to work in the thickets and if you do not burn before planting. Farming operation and activities become easier if we burn. Even now that we are aware of the need not to use slash and burn because you destroy the organisms that improve the soil, we do burn the fresh weeds on the forest land" (FGD with farmers of Wlewlexena Kododo Dededa group in Kobo No. 2 in the Jasikan district)

7.6 Adoption of ISFM

Since ISFM is a suite of different agricultural practices, we measure the impact of Africare's intervention on each of these individual practices, namely

1) Ploughing
2) Application of manure
3) Application of other organic fertilizers
4) Growing leguminous trees
5) Growing leguminous plants (cowpea)
6) Growing leguminous plants (groundnut)
7) Using purchased seed
8) Using weedicide
9) Using pesticide
10) Value of NPK used

To measure the impact of Africare's intervention on adoption of ISFM, we estimate DID models for each of the practices above, with and without matching. The first nine measurements are indicator variables, and the value of NPK used is measured as the total value of NPK used divided by crop area.

Our estimates indicate no significant difference in adoption of ISFM practices between treatment and control from baseline to endline, with the exception of a 7-percentage point increase in pesticide use. Results are shown in the table 13.

Results from the qualitative study

The results from the mid-term qualitative study shed light on why adoption rates for ISFM technologies and practices in the project area were low. This section includes some key findings on adoption of ISFM practices from the qualitative study. The study found that while farmers are generally aware of methods to detect low soil fertility, they are not adopting these practices. Farmers are aware of the benefits of applying inorganic fertilizers and some have adopted this practice on their farms. Some of the primary barriers to adoption of inorganic fertilizers include – high costs, access to fertilizers, changing attitudes, and bad experiences with weather. Green manuring, farmyard manuring, composting, and cover cropping technologies have largely not been adopted by farmers because of limited knowledge and skills and their incompatibility with farmers' current practices. Smallholder farmers are aware that growing legumes (crop rotation with maize and cassava) can improve soil fertility but few farmers have adopted these practices because this is not commercially viable. The smallholder farmers interviewed were divided on the practice of leaving the crop residues on the field as some use tractors for ploughing and zero tillage. The study found that farmers do not deliberately mulch their farms. While farmers were aware of the consequences of burning the vegetation to discourage the presence of unwanted organisms, this practice is not entirely in their control because hunters and headmen deliberately burn their farms.

Smallholder farmers report having adopted herbicides to control weeds because AEAs have taught them how, and because farmers reported that it is not a costly practice. They have also adopted careful weed management to obtain higher yields. Farmers have adopted improved seeds introduced by this project to some extent. However, farmers still use local seeds due to the high costs of improved seeds and non-availability. In addition, farmers practice the recommended rate of two seeds per hill and line planting because they have observed the benefits of this practice. The variations in the spacing between and within plants are primarily due to labour cost and availability. Quotes from the field below reflect these findings:

> "*We did not use fertilizer prior to introduction of the project. Rather, it is the project that introduced us to use of fertilizer"* (from an FGD in Afadjato South, North Dayi, and Jasikan).
>
> *"I think it is the use of local that is resulted in the poor yield. Prior to the project that I did not use improved seeds, I got only 5 bags (600kg) for an acre (0.41ha) for maize and 2 bags (240kg) for an acre (0.41ha) for cowpea"* (FGD with members of Biakoye Farmers Association in Atonkor in the Jasikan district).

7.7 Labour and yield

We measure the impacts of Africare's intervention on labour by looking at total man-days divided by crop area, and total woman-days divided by crop area for maize in the major season. For yield, we use self-reported harvest quantity and cultivated area. We do not find a significant impact on any of these three outcome variables using a DID estimator, with and without matching. Results are shown in the Table 13.

7.8 Cost effectiveness analysis

In addition to collection of data from treatment and control households, data were collected from AEAs, Africare, and MoFA on the costs of project implementation. These data can be employed to analyse the cost-effectiveness of Africare's project (and AGRA's investment in the project) and, with the use of several additional assumptions, the cost-effectiveness of an out-of-sample scaling-up of Africare's project.

Key elements of cost-effectiveness analysis include (1) accurately defining and measuring the direct costs of the project, (2) accurately identifying and measuring indirect costs (e.g., tier-specific costs of implementation incurred at the community, district, regional, and central levels), (3) capturing the time dimension of the project (periodicity and the manner in which capital costs are annualized), (4) disaggregating project components and activities, (5) establishing a credible comparison between costs associated with the non-treated counterfactual, and (6) introducing credible assumptions for out-of-sample scale-up analysis. These issues are typically addressed in cost-effectiveness analyses of health and education projects (see reviews by Galárraga et al. (2009) and McEwan (2012)), but rarely applied to agricultural extension projects such as the subject of this evaluation. See Mogues, Mueller, and Kondylis (2017) for an application to an extension project in Mozambique.

Given the results presented above – insignificant effects of the treatment on awareness, knowledge, or adoption – it is difficult to establish any outcomes against which costs can be detailed. As such, we choose not to proceed with the cost-effectiveness analysis at this juncture.

8. Discussion

The impact evaluation methods used in this study were selected to ensure that farmers targeted by or participating in the project intervention could be credibly compared to similar households based on observable characteristics, and that the outcomes observed could be credibly compared both before and after the intervention to account for any exogenous changes affecting all farmers sampled in the study. This is, of course, a second-best approach to evaluation because it does not account for differences in unobservable characteristics that might affect the outcomes of interest.

In the absence of randomized assignment of Africare's intervention, the DID approach with matching techniques was considered the next-best option. As a first step in this direction, we explored results that only draw on the DID estimations without matching. This is done because changes in treatment assignment during project implementation may threaten our ability to identify an area of common support through matching.

Of course, this DID approach carries with it a set of concerns for internal validity. Embedded in the DID principle are assumptions that (1) the allocation of intervention was not determined by outcome, or in other words, if a farmer in the control group had been placed in the treated group, that farmer would have experienced outcomes similar to a similar farmer in the treated group, all else being equal; (2) the treated and control farmers have parallel trends in outcomes, or in other words, in the absence of the intervention, the differences in the two groups would remain constant over time; (3) the treatment and control groups remained stable; and (4) there were no spill over effects, or in other words, the outcomes for each individual or household are independent of the treatment assignment of other individuals or households. The last two assumptions are part of the so-called stable unit treatment value assumption (SUTVA) and are likely to be violated when there are interactions between the treated and control groups (Rubin 1977; Lechner 2011). Below we consider each of the potential threats to internal validity from violations of these four assumptions.

We argued earlier that the external validity of this evaluation is fairly limited given the socioeconomic and agro-ecological diversity of Ghana. As such, the evaluation is likely valid only to districts and populations in the Volta Region that are most similar to Africare's project site. We also argued earlier that the treatment design itself was unlikely to have had an impact on the behaviour of either the treated or control farmers (i.e., reactivity effects).

But in terms of the assumptions about parallel trends in outcomes, the crucial question revolves around how Africare allocated its treatment or implemented its project. The initial assumption employed in this evaluation was that control farmers in control districts could be selected based on observable characteristics such that any outcomes observed for a control farmer, had they been in the treated group, would have been the same as outcomes observed for a similar farmer in the treated group. Careful consideration was given to the selection of control district by all parties involved in the project and its evaluation. While there is no statistical test for the parallel trends assumption, the historical data used as a guide for selecting control districts, and the fact that the intervention was evaluated over a relatively short time (less than two years) gives confidence that the assumptions were not violated.

Furthermore, the assumption about treatment and control group stability was likely violated during the evaluation. The mid-course shift in the project's targeting strategy – from a focus on FBOs and their members, to a less targeted approach designed to recruit *any* farmers and communities that demanded their services – had profound effects on whether treated households that were surveyed at baseline ever received the treatment.

9. Specific findings for policy and practice

The key finding emerging from this evaluation is the dual importance of carefully designing a project with evaluation in mind, while also designing an evaluation with a project in mind. In this particular case – and despite best efforts made on the part of the implementing partners, evaluators, and their respective donors – the design the project and its evaluation were not conducted in a sufficiently coordinated manner.

As a result, little is ultimately known about the impact of Africare's project. Certainly, the issue that Africare sought to address – the negative yield and income effects of soil degradation in smallholder production systems – is important to food security, growth, and development in the Volta Region and in Ghana more generally. It is possible that the project had an impact on the targeted households that simply could not be captured by the measurements used here. It is also possible that the project had little or no impact as the analysis set forth above suggests.

Below, we summarize the lessons learned as they pertain to the different audiences.

9.1 Donors and evaluators

Unfortunately, several confounding factors emerged during implementation. Specifically, the project's initial strategy to reach its planned targets was likely over-ambitious, as were its targets for outcomes (e.g., changes in yield and income). Mid-course adjustments were necessary to keep the project on track, and the implementing partner cannot be faulted for making such adjustments. However, the adjustments do militate against good evaluation design and execution. This suggests the need for more careful and measured approaches to the funding and deployment of evaluations such as the one conducted here.

An important corollary of this finding is that rigorous evaluations themselves may be appropriate only at certain project sizes, scales, and durations. Africare's project targeted a fairly small population within a short timeframe to achieve outcomes that can take years to accumulate and observe. The very idea that a project such as this can be the subject of rigorous evaluation may have been itself too ambitious.

Specifically, when identifying a project for evaluation, donors might ask the following questions:

1. Does the project design lend itself to rigorous evaluation?
2. Does the project intervention timeline allow for an evaluation to be embedded in the design?
3. Does the implementing agency have a learning objective? Are they interested in the results of an evaluation of their project?
4. Does the implementing agency have the capacity to understand and undertake a collaboration with the evaluating agency?

If the answer to any of these questions is no, the decision to conduct an evaluation may warrant further consideration.

An additional consideration is the cost of the evaluation (the "value for money" question). Ideally, a comprehensive evaluation should be relatively small in comparison with an overall project budget. Our past experience suggests that evaluation budgets might optimally reach approximately 10 percent of the project budget, and certainly no greater than the project budget itself. If the cost of an evaluation is too large relative the project budget, then the decision to conduct an evaluation may again warrant further consideration.

9.2 Implementers

In this specific project, the implementing agency corrected its course of action under the project to meet its targets. Its decision to do so was probably well taken given its year-on-year performance. This raises several key lessons for implementing agencies participating in an impact evaluation to consider more closely, as follows.:

1. Invest time to identify one's learning objectives from a project, understand the objectives of the evaluation, and consider how it can be designed to help improve programming.
2. To get the greatest benefit from an evaluation, work closely with the evaluators and keep clear communications channels to ensure they are kept updated of any changes in the implementation, however small these changes might seem.
3. Maintain comprehensive monitoring and evaluation data to inform the impact evaluation at all stages of the process.

9.3 Evaluators

In this section, we reflect on the 'best possible case' outcome and steps evaluators can take to ensure this outcome is reached. For an evaluator, the best possible case would obviously be a situation where project donors and project implementers have carefully considered the questions and issues identified above and resolve to execute the project with an impact evaluation fully embedded in the design and implementation of the project. The remaining challenges—whether the evaluation is experimental in design and whether treatments can be randomly assigned— are then merely technical.

However, since in practice, project implementers are likely to change the project along the way in response to internal and external factors that may or may not be within their control, implementation changes are likely to occur that adversely affect evaluation goals. This was our experience with the subject project, so we offer several steps that an evaluator can take before project and evaluation inception to reduce or mitigate the impacts of changes in implementation.

First, consider that the main task for an evaluator is to assess the risks and threats to internal validity of an evaluation's findings before project implementation and, with assistance from donors and implementers, assess the different paths along which the project might unfold on the ground. Evaluators do so bearing in mind that development projects often have multiple objectives that might be in conflict with the goals of an evaluation. This task is more complicated than it may seem at first glance, because it requires evaluators (who are often researchers) to put themselves in the position of project implementers, model future scenarios and possible responses, and assess how the responses might affect the rigor of the evaluation.

From our experience, three areas deserve special attention from evaluators when considering ways to remove threats to internal validity arising from project implementation. First, it is important to arrive at a shared agreement with project implementers on the counterfactual for the evaluation. Evaluators are trained to search for credible counterfactuals (i.e., situations in which a population experiences a state of the world that would have happened in the absence of the intervention, as compared to the observed situation) to ensure the internal validity of impact estimates. This often necessitates identifying control groups outside of the project area, as was the case for our evaluation. On the other hand, implementers may be more concerned about including groups in the control group who actually reside in or are closely linked to the intervention area, since these are often the individuals they target in project expansion or scale-up activities, and since implementers often want useful baseline information on these groups and individuals. Unfortunately, these same groups and individuals may be more likely to be treated while in the control group (contamination) or benefit indirectly from the treatment (spillovers), making them poor candidates for the control group and the construction of a credible counterfactual.

A second and related point is the need for evaluators to determine how a project can adhere strictly to definitions of treatment and control groups for the purposes of both project implementation and evaluation. Defining the evaluation counterfactual leads to the next logical step of defining the treatment and control groups for the evaluation. Changing the composition of the groups after implementation commences, as occurred in this project for the treatment group, compromises the effectiveness of the evaluation.

A third area that deserves attention from evaluators is the collection and use of project monitoring data. For many good reasons, data collection for monitoring purposes is often undertaken separately from data collection for impact evaluation purposes. One of the lessons learned from this project is that it may be helpful to reconsider this division of activities in some cases. Project monitoring data can augment the information gathered from more comprehensive baseline and endline surveys, because monitoring data are better, in some cases, for tracking deviations from planned implementation arrangements. For this evaluation, we have benefited from the use of project monitoring data to re-classify the treatment and control groups for some of the analyses presented in this report. This type of *ex post* arrangement could be improved by, for example, strengthening collection of monitoring data in selected areas, where evaluation data has already been collected, and by matching the units, variables, and identification numbers for collecting monitoring data to those used for collecting evaluation (baseline) data.

In summary, there are several strategies to address the discontinuities between project implementation and evaluation. All of them rest on close coordination between implementer and evaluator, agreement on learning objectives of mutual interest, early cooperation in design processes for both the project and evaluation, and careful accommodation of defining counterfactuals, maintaining treatment and control groups, and managing data collection, use, and analysis. Many of these elements did not come together in the subject evaluation despite best efforts by all parties. Nonetheless, the experience gained for the donors, implementers, and evaluators involved in this process will likely help advance cooperation and coordination around future project investments and evaluation designs.

Online appendixes

Online appendix A – Questionnaires

http://www.3ieimpact.org/sites/default/files/TW4.1022-Online%20appendix%20A-Questionnaires.pdf

Online appendix B – Sample design

http://www.3ieimpact.org/sites/default/files/TW4.1022-Online%20appendix%20B%20-%20Sample%20design.pdf

Online appendix C – Pre-analysis plan

http://www.3ieimpact.org/sites/default/files/TW4.1022-Online%20appendix%20C%20-%20Pre-analysis%20plan.pdf

Online appendix D – Power calculations

http://www.3ieimpact.org/sites/default/files/TW4.1022-Online%20appendix%20D%20-%20Power%20calculations.pdf

Online appendix E – Monitoring plan

http://www.3ieimpact.org/sites/default/files/TW4.1022-Online-appendix-E-Monitoring-plan.xlsx

Online appendix F – Descriptive statistics

http://www.3ieimpact.org/sites/default/files/TW4.1022-Online-appendix-F-Descriptive-statistics.pdf

Online appendix G – Content knowledge test

http://www.3ieimpact.org/sites/default/files/TW4.1022-Online-appendix-G-Content-knowledge-test.pdf

Online appendix H – Avoiding bias

http://www.3ieimpact.org/sites/default/files/TW4.1022-Online-appendix-H-Avoiding-bias.pdf

References

Abadie, A, and Imbens, GW 2016, 'Matching on the estimated propensity score', *Econometrica*, 84 (2), pp.781–807.

Africare 2015, *Grant proposal to the Alliance for a Green Revolution in Africa (AGRA): Scaling-out of integrated soil fertility management technologies in the Volta Region, Ghana,* Africare, Washington, DC and Accra.

Anderson, JR 2007. 'Agricultural advisory services' Background paper for World Development Report 2008, Agriculture for Development, World Bank, Washington, DC.

Birner, R, Davis, K, Pender, J, Nkonya, E, Anandajayasekeram, P, Ekboir, J, Mbabu, A, Spielman, D, Horna, D and Benin, S 2009, 'From best practice to best fit: A framework for analysing agricultural advisory services worldwide', *Journal of Agricultural Extension and Education* 15 (4): 341–355.

Bonan, J and Pagani, L 2017, 'Junior farmer field schools, agricultural knowledge and spillover effects: Quasi-experimental evidence from Northern Uganda', *Journal of Development Studies,* pp. 1–16.

Bulte, E, Beekman, G, Di Falco, S, Hella, J and Lei, P 2014, 'Behavioral responses and the impact of new agricultural technologies: Evidence from a double-blind field experiment in Tanzania', *American Journal of Agricultural Economics*, 96 (3), pp.813–830.

Conley, T and Udry, C 2001, 'Social learning through networks: The adoption of new agricultural technologies in Ghana,' *American Journal of Agricultural Economics*, 83 (3), pp.668–673.

Cole, S, Giné, X, Tobacman, J, Topalova, P, Townsend, R and Vickery, J 2013, 'Barriers to household risk management: Evidence from India', *American Economic Journal: Applied Economics*, 5 (1), pp.104–135.

Cooke, E, Hague, S, and McKay, A 2016, *The Ghana poverty and inequality report: Using the 6th Ghana Living Standards Survey,* University of Sussex.

Davis, K 2008, 'Extension in sub-Saharan Africa: Overview and assessment of past and current models and future prospects', *Journal of International Agricultural and Extension Education*, 15 (3), pp.15–28.

Dercon, S and Christiaensen, L 2011, 'Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia', *Journal of Development Economics*, 96 (2), pp.159–173.

Davis, K, Nkonya, E, Kato, E, Mekonnen, DA, Odendo, M, Miiro, R and Nkuba, J 2012, Impact of farmer field schools on agricultural productivity and poverty in East Africa', *World Development*, 40 (2), pp.402–413.

Diao, X, Hazell, P, Kolavalli, S and Resnick, D 2018, *Why the lion hasn't roared louder: Ghana's agricultural and economic transformation,* draft manuscript.

Fairhurst, T (ed) 2012, *Africa soil health consortium: Handbook for integrated soil fertility management,* CABI, Nairobi.

Feder, G., 1980. Farm size, risk aversion and the adoption of new technology under uncertainty. *Oxford Economic Papers*, *32*(2), pp.263-283.

Feder, G, Just, RE and Zilberman, D 1985, 'Adoption of agricultural innovations in developing countries: A survey', *Economic Development and Cultural Change*, 33 (2), pp.255–298.

Foster, AD and Rosenzweig, MR 2010, 'Microeconomics of technology adoption', *Annual Review Of Economics*, 2 (1), pp.395–424.

Foster, AD and Rosenzweig, MR 1995, 'Learning by doing and learning from others: Human capital and technical change in agriculture', *Journal of Political Economy*, 103 (6), pp.1176–1209.

Galárraga, O, Colchero, MA, Wamai, RG, and Bertozzi, SM 2009, 'HIV prevention cost-effectiveness: A systematic review', *BMC Public Health* 9 (S1), p. S5.

Giné, X, Townsend, R and Vickery, J 2008, 'Patterns of rainfall insurance participation in rural India', *The World Bank Economic Review*, 22 (3), pp.539–566.

Griliches, Z 1957, 'Hybrid corn: An exploration in the economics of technological change' *Econometrica,* 25 (4), pp. 501–522.

Hanna, R, Mullainathan, S and Schwartzstein, J 2014, 'Learning through noticing: Theory and evidence from a field experiment', *The Quarterly Journal of Economics*, 129 (3), pp.1311–1353.

Hill, RV, Hoddinott, J and Kumar, N 2013, 'Adoption of weather-index insurance: Learning from willingness to pay among a panel of households in rural Ethiopia', *Agricultural Economics*, 44 (4–5), pp.385–398.

Haug, R 1999, 'From integrated rural development to sustainable livelihoods: What is the role of food and agriculture?', *Forum for Development Studies* 26 (2), pp. 181–201.

Ho, D, Imai, K, King, G, and Stuart, E 2007, 'Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference.' *Political Analysis*, 15, pp. 199–236.

Imbens and Wooldridge, 2007, 'Lecture notes: Difference-in-differences estimation', NBER http://www.nber.org/WNE/lect_10_diffindiffs.pdf.

Jack, BK 2013, 'Market inefficiencies and the adoption of agricultural technologies in developing countries' CEGA Working Papers, University of California Berkeley.

Klerkx, L, and Leeuwis, C 2008, 'Matching demand and supply in the agricultural knowledge infrastructure: Experiences with innovation intermediaries', *Food Policy*, 33 (3), pp. 260–276.

Lechner, M, 2011. 'The estimation of causal effects by difference-in-difference methods', *Foundations and Trends in Econometrics*, 4 (3), pp.165–224.

Kondylis, F, Mueller, V and Zhu, S 2014, 'Seeing is believing? Evidence from an extension network experiment', *World Bank Policy Research Working Paper* 7000, World Bank, Washington, DC.

McEwan, PJ 2012, 'Cost-effectiveness analysis of education and health interventions in developing countries', *Journal of Development Effectiveness* 4 (2), pp. 189–213.

Mogues, T, Mueller, V, and Kondylis, F 2017, 'Cost-effectiveness of community-based gendered advisory services to farmers: Analysis in Mozambique and Tanzania', IFPRI Discussion Paper 1613, International Food Policy Research Institute, Washington, DC.

Molini, V and Paci, P 2015, *Poverty reduction in Ghana: Progress and challenges.* Washington, DC: World Bank.

Munshi, K 2004, 'Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution,' *Journal of Development Economics*, 73 (1), pp.185–213.

Nkonya E, Place, F, Kato, E and Mwanjololo, M 2015, 'Climate risk management through sustainable land management in sub-Saharan Africa', in R Lal, B Singh, D Mwaseba, D Kraybill, D Hansen and L. Eik, (eds), *Sustainable intensification to advance food security and enhance climate resilience in Africa,* Springer, Switzerland.

Place, F, Barrett, C B, Freeman, HA, Ramisch, JJ and Vanlauwe, B 2003, 'Prospects for integrated soil fertility management using organic and inorganic inputs: Evidence from smallholder African agricultural systems', *Food Policy*, 28 (4), pp. 365–378.

Rogers, EM 1962, *Diffusion of innovations,* 1st edn, The Free Press, New York.

Rubin, DB 1977, 'Assignment to treatment group on the basis of a covariate', *Journal of Educational Statistics*, 2 (1), pp.1–26.

Schultz, TW 1975, 'The value of the ability to deal with disequilibria', *Journal of Economic Literature*, 13 (3), pp. 827–846.

Vanlauwe, B 2004, *Integrated soil fertility management research at TSBF: The framework, the principles, and their application*, Academy Science Publishers, Nairobi.

Vanlauwe, B, Diels, J, Sanginga, N and Merckx, R 2005, 'Long-term integrated soil fertility management in South-western Nigeria: Crop performance and impact on the soil fertility status', *Plant and Soil,* 273 (1-2), pp. 337