The Effects of Decentralized and Video-based Extension on the Adoption of Integrated Soil Fertility Management – Experimental Evidence from Ethiopia

Denise Hörner[a,*], Adrien Bouguen[b], Markus Frölich[c], Meike Wollni[a]

[a] Department of Agricultural Economics and Rural Development, University of Goettingen/Germany

[b] Agriculture and Resources Economics Department, University of California, Berkeley

[c] University of Mannheim/Germany

* Corresponding author. E-mail: denise.hoerner@uni-goettingen.de

Abstract

The slow adoption of new agricultural technologies is an important factor in explaining persistent productivity deficits among smallholders in Sub-Saharan Africa (SSA). Farmers in particular delay the uptake of technology packages, i.e. practices that should be applied in combination. Since lagged adoption is often attributed to knowledge gaps, the effectiveness of extension services as key information delivery mechanisms is crucial, but largely understudied. In recent decades, extension systems in many SSA countries are undergoing substantial transformations, away from "top-down" towards more inclusive "bottom-up" models that involve farmers as active stakeholders in the technology transfer process. In these decentralized and participatory approaches, only few model farmers are trained directly by extension agents and should then pass on knowledge to other farmers, often via group-based learning. From there, information should ideally trickle down to all other households in a community. Yet, since only a share of farmers actively participates in extension activities, the success of this approach heavily relies on information spillovers between farmers. In light of selective attention theory, learning and teaching failures might lead to incomplete information transmission, while "reminders" of potentially neglected knowledge dimensions can counteract information failures. In this study we assess the effects of a decentralized extension program and an additional video intervention on the adoption of *integrated soil fertility management* (ISFM) among small-scale farmers in Ethiopia using a randomized control trial. ISFM is widely promoted in SSA as a promising concept to enhance soil fertility and productivity by using organic and inorganic soil amendments simultaneously. We find that both extension-only and extension combined with video induce ISFM adoption as well as gains in knowledge. We further find evidence for spillovers to farmers in treatment communities that do not actively participate in extension activities regarding increased adoption of ISFM practices at the household level. Yet, when it comes to the *integrated* use of the practices on the same plot, extension alone does not significantly influence these non-actively involved farmers, while the video intervention shows a significant complementary effect for this group. A causal mediation analysis further reveals that increases in knowledge explain part of the treatment effects on adoption.

Acknowledgement

This research was financially supported by the German Research Foundation (DFG) through grant number RTG1666 (GlobalFood) and project number 390367541, as well as the German Agency for International Cooperation (GIZ). The authors thank all members of the Integrated Soil Fertility Management Project (ISFM+) for their invaluable support, and Ibrahim Worku for his excellent field work coordination as well as the whole team of research assistants and enumerators.

1. Introduction

The slow adoption of new agricultural technologies is an important factor in explaining persistent productivity deficits among smallholder farmers, especially in Sub-Saharan Africa (SSA). Moreover, underinvestment in the productive capacity of their soils is viewed as a major cause for self-reinforcing poverty traps for the rural poor (Barrett and Bevis, 2015). Recent evidence shows that farmers in particular delay the uptake of system technologies, i.e. packages of agricultural practices that should be applied in combination in order to deploy their full productivity-enhancing potential (Noltze et al., 2012; Sheahan and Barrett, 2017; Ward et al., 2018). Even though farmers might simultaneously adopt several components at the household level, it is a common behavior to scatter them across plots rather than combining them on the same plot, leaving much of their synergistic potential unexploited (Sheahan and Barrett, 2017).

A frequently cited barrier to the adoption of agricultural innovations is the lack of information and knowledge (Foster and Rosenzweig, 1995; Aker, 2011). Providing agricultural extension services to farmers can bridge the knowledge and capacity gap, as its overarching goal is knowledge transfer from researchers to farmers. Extension activities often aim at providing awareness for improved agricultural practices and instructions on *how* to implement them, but frequently disregard the importance of providing sufficient information on *why* certain practices are beneficial (Rogers, 1995; Anderson and Feder, 2007). Yet, individuals' "competence to decide whether or not to adopt" a technology can be facilitated by being well informed about their underlying principles and mechanisms due to enhanced capacity of appraising the consequences of adoption (Rogers, 1995: 166).

Extension systems are frequently subject to a series of shortcomings, such as high bureaucratic burden, excessive costs of direct trainings, limited geographic coverage, and exclusion of marginalized, resource-poor farmers (Anderson and Feder, 2007; Aker, 2011). In recent decades, this gave rise to the introduction of decentralized approaches, especially in SSA, where extension agents only train selected farmers (often referred to as contact, lead or model farmers) to adopt new techniques on their farms. These model farmers should then in turn train other farmers, often organized in farmer groups, in a participatory and experiential way. This goes along with a shift in perspective from a "top-down" to a more inclusive "bottom-up" strategy by involving farmers as active stakeholders in the technology transfer process. Eventually, exposure to on-farm demonstrations, trained contact farmers and group members is supposed to spur broader adoption in a community through peer learning (Gautam, 2000;

Swanson, 2008; Kondylis et al., 2017). Oftentimes, these developments are accompanied with a change from a mere output-growth perspective to a more holistic view, promoting sets of technologies that achieve productivity increases and sustainable use of natural resources at the same time (Swanson, 2008).

A crucial question with respect to decentralized extension models is how effectively information is passed on from farmer to farmer. Some studies suggest positive effects of participatory extension approaches like farmer field schools on participants' knowledge, adoption of new practices or income (Feder et al., 2004; Godlandt, 2004; Davis et al, 2012). Yet, multiple studies conclude that knowledge gains hardly trickle down to neighboring farmers (Rola et al., 2002; Feder et al., 2004; Tripp et al., 2005), and that increased technology adoption among trained farmers does little to change behavior among their peers (Berg and Jiggins, 2007; Kondylis et al., 2017), or that diffusion highly depends on similarity between communicators and target farmers (BenYishay and Mobarak, 2018) as well as other context-specific forms of social capital prevalent in communities (Pamuk et al., 2014).

In their study on pit-planting knowledge transmission from extension agents to lead farmers to other farmers, Niu and Ragasa (2018) show that information loss takes place due to both teaching and learning failures. They find that even though knowledge is transmitted, important dimensions get lost along the chain due to selective attention on both the side of the information senders and receivers. Due to the mental costs attached to processing new information, individuals seem to neglect information they do consider less important. Yet, literature suggests that reminders about knowledge dimensions commonly neglected by farmers can help to offset teaching and learning failures (Hanna et al., 2014; Niu and Ragasa, 2018).

In this research, we assess the effect of a decentralized, participatory extension model as well as an additional video intervention on the adoption of and knowledge about a package of *integrated soil fertility management* (ISFM) practices among small-scale subsistence farmers in rural Ethiopia. We use a randomized control trial (RCT) with two treatment arms, extension-only and extension in combination with video, as well as a control group. In particular, we are interested in whether exposing farmers to explanations on *why* all individual ISFM components are important via video can counterbalance incomplete information diffusion that might occur in a decentralized extension set-up and therefore, foster adoption of the practices and the ISFM technology package as a whole.

ISFM is widely promoted in SSA, since it is viewed as a strategy to sustainably intensify agricultural productivity and combating land degradation, caused by excessive deforestation and inappropriate agricultural land use practices, such as overgrazing, improper crop rotations, insufficient fallow periods, intensive tillage or immoderate use of mineral fertilizers (Barrow, 1991). A fundamental feature of the ISFM paradigm is the *integrated* use of inorganic and organic soil amendments (such as compost, crop residues or manure) in order to enhance both nutrient availability as well as the soil's capacity to absorb nutrients, and is ideally complemented by the use of improved seeds. Complementarily, ISFM aims at a general improvement of agronomic techniques, like line application of seeds and fertilizers, and the introduction of practices such as crop rotation, cereal-legume intercropping, minimum tillage, mulching, residue retention, agroforestry or green manuring, targeted to local conditions (Place et al., 2003; Vanlauwe et al., 2010). ISFM provides a promising concept to tackle soil degradation and increase productivity among smallholders in developing countries. Yet, knowledge constraints are likely to hinder its uptake or might lead to incomplete adoption of this relatively complex system technology. This is even more likely since the adoption of technology packages is typically not viewed as one single, but rather as a series of separate (even if interrelated) decisions (Ward et al., 2018), so that farmers might only take up those practices they consider more important or that are easier to learn. In addition, some practices involve additional costs (e.g. for improved seeds or mineral fertilizer) or labor (e.g. for seeding in lines or compost production), as well as trade-offs with other livelihood activities (animal manure as fuel, crop residues for animal feed).

The study is conducted within the Integrated Soil Fertility Management Project of the German Agency for International Cooperation in Ethiopia and focuses on a package of five ISFM technologies: Compost, blended fertilizer, improved seeds, line seeding and liming for acidic soils. The extension model builds on model farmers as entry points for ISFM knowledge, who should pass on their knowledge to members of so-called Farmer Research and Extension Groups, from where information should then ideally trickle down to other farmers in a community. Since farmers might not consider each ISFM component as equally important or some pieces of information are more difficult to grasp, knowledge transmission is likely to occur incomplete. To counteract this, we designed a second intervention in form of a video emphasizing the importance of perceiving ISFM as a holistic concept and providing information on the underlying principles of each component. Previous research has shown that video as information delivery channel has the potential to induce behavioral changes in farming communities (Van Mele, 2006; Zoussou et al., 2010; Bernard et al., 2014), can increase the effectiveness of standard extension activities (Gandhi et al., 2009; Van Campenhout et al., 2017; Vasilaky et al., 2018) and even trigger knowledge increases in areas not explicitly mentioned in the videos (Van Campenhout et al., 2017).

Despite the high policy-relevance, there is little rigorous evidence on the impact of different extension models on the uptake of agricultural innovations (Kondylis et al., 2017; BenYishay and Mobarak, 2018; De Brauw et al., 2018; Ogutu et al., 2018). We seek to contribute to the literature on the impact of decentralized extension and agricultural learning by using an experimental research design. We further assess the differential effects of the extension model on those who actively take part in the activities and those who at most benefit indirectly, in order to assess the effectiveness of information spillovers as a key principle of decentralized extension. Moreover, whereas previous studies mostly focus on the adoption of individual technologies, or sets of technologies that are viewed as a series of binary adoption decisions, we also study the integrated uptake of ISFM practices. Additionally, we examine the importance of gains in ISFM knowledge, in particular of knowing *why* a technology works, as a driver for adoption using a causal mediation analysis. And ultimately, we contribute to the emerging literature on the use of media and other non-traditional ways of agricultural education (Aker, 2011).

The remainder of this article is structured as follows: In the next section we provide an outline of the context and the conceptual model underlying our study. Subsequently, we describe the experimental design, the data used for analysis, as well as our estimation strategy. In the results section, we first assess the impact of the interventions on ISFM adoption, before analyzing treatment effects on knowledge as potential impact pathway. The last chapter discusses implications of our findings and concludes.

2. Setting and Conceptual Framework

Agriculture forms the backbone of Ethiopia's economy, since it accounts for over 35% of the country's GDP and presents the major income source for around three-fourths of the population (CIA, 2018). Five cereals – teff, maize, wheat, barley and sorghum – are the most important staple food crops, both in terms of production and consumption (Taffesse et al., 2011). Despite the importance of the sector and substantial output growth in recent years, agricultural yields remain comparatively low with average cereal yields below 2.5 t/ha (FAO, 2016). Land degradation and declining soil fertility are among the most serious problems for the Ethiopian agriculture. In the past decade, the Ethiopian government has responded to these challenges with considerable investments in the extension system, estimated to around 2% of the agricultural GDP (Spielman et al., 2010). At the same time, rural advisory services have undergone substantial structural changes, away from a centralized top-down approach, typically only reaching few, rather resource-rich farmers, towards a more decentralized inclusive outreach program (Belay, 2003). In 2010, the Ethiopian government announced a “participatory extension system” as core of the national strategy for agricultural growth and development. Cornerstone of this strategy is a strong emphasis on grassroots farmer groups and strengthening of peer-to-peer learning (ATA, 2014).

2.1 Study Context

Our study is conducted within the Integrated Soil Fertility Management Project (ISFM+ project) of the German Agency for International Cooperation (GIZ), which was launched in mid-2015 and operates in the three Ethiopian highland regions Amhara, Oromia and Tigray. The ISFM+ project is a component of the broader Ethiopian Sustainable Land Management Programme (SLMP)[1] and only operates in districts (in Ethiopia called Woredas) where physical land rehabilitation measures (stabilization of hillsides, erosion control measures) have been successfully introduced by the SLMP. During the initial phase of the project from 2015 to 2018, the use of five so-called “quickwin technologies” was promoted among small-scale farmers in the three regions, since the combination of these practices is considered to boost on-farm biomass production within a relatively short period of time. The quickwin package consists of the following practices: *Compost*, prepared of crop residues or other plant materials and animal dung, is supposed to increase soil organic matter, thus improving nutrient supply, soil biota as well as water holding capacity. *Blended fertilizer* refers to inorganic fertilizers that are aligned to a specific location’s soil type and therefore provide a balanced nutrient supply. It is commonly composed of nitrogen (N), phosphor (P), potassium (K), sulfur (S), zinc (Zn) and boron (B) and should replace the widely used DAP fertilizer. Fertilizer blending factories have recently started their operation in Ethiopia. *Improved seeds* should increase biomass production of both grain and residues and are distributed to model farmers by the project for all major crop types. *Line seeding* is promoted to replace the common practice of broadcasting seeds. It is thought to reduce competition for space, nutrients and water among plants and thus, lead to more vigorous growth. At the same time, line seeding allows to target inputs directly to the plants, and hence reduce required amounts and enhance efficiency. *Lime* application is promoted in regions where soil suffers from acidity in order to normalize its pH value.

---

[1] Beginning of 2018, the SLMP has been replaced by the successor project named Sustainable Use of Rehabilitated Land for Economic Development (SURED).

In line with the country's national strategy, it was chosen to use a decentralized "participatory learning and extension approach" as delivery mechanism of the promoted practices (described in more detail in section 3). The ISFM+ provides capacity building measures among government agricultural advisory staff, which includes experts from provincial and Woreda Bureaus of Agriculture as well as development agents, the extension officers at Kebele level, (the lowest administrative unit in Ethiopia) who should then train model farmers on ISFM. Model farmers act as leaders of so-called Farmer Research and Extension Groups (FREGs) and should pass on their ISFM knowledge to the group members.

2.2 Conceptual Model

As pointed out earlier, the key feature of ISFM is the integrated use of a range of different practices. Hence, it is pivotal for farmers to learn about each of its individual components as well as the necessity of applying them jointly. This is, however, frequently neglected by farmers, which may be a result of learning gaps.

A useful way of conceptualizing this shortcoming is to view it as a learning failure in the framework of selective attention theory, which is borrowed from psychology (Schwartzstein 2014), and has been applied in the context of agricultural learning before. Hanna et al. (2014) as well as Niu and Ragasa (2018) developed a set of assumptions based on Schwartzstein's (2014) model of selective attention that are relevant to the agricultural technology adoption context: First, a new technology comes along with a certain set of parameters that are unknown and must be learned by a farmer, e.g. through trainings, visits, or farmer-to-farmer extension. Yet, farmers do often not consider all aspects equally important and therefore, a priori, attach different weights to these. Second, paying attention involves costs, because learning requires capacities in the form of mental energy and time, and individuals need to economize these resources. Third, farmers seek to maximize their net payoffs, resulting from expected yields minus attentional costs. Consequently, even when full information on a new technology is readily available through trainings, field demonstrations or on neighbors' fields, farmers may not be able to pay attention to each of its parameters due to resource boundaries, and therefore need to decide which dimensions to focus on.

In the case of a system technology that requires learning about several individual practices, a resource-constrained farmer might – consciously or unconsciously – base the decision which components to focus on not only on how important she or he considers a certain practice, but also on its level of complexity. Since learning more complex technologies requires more cognitive energy, payoff-maximizing farmers will only learn them when they are sufficiently convinced of their benefits, but otherwise disregard.

Knowledge dimensions that have been neglected from the beginning are often continuously ignored throughout the further process of experimentation and implementation, simply because farmers did initially not pay attention to them, due to low perceived importance or high perceived complexity. In that sense, a learning failure essentially stems from a failure to notice (Niu and Ragasa, 2018). As a result, farmers may persistently stick to suboptimal choices or applications of technologies, if they do not get reminded of the ignored parts. Conversely, reminders of neglected dimensions of a technology (package) may help to overcome this learning failure and alter farmers' behavior (Hanna et al., 2014; Niu and Ragasa, 2018).

The ISFM technology package promoted via the above described extension intervention consists of several individual components. Yet, due to a failure to notice the importance of each individual – and in particular the more complex – components, we expect learning and teaching along the knowledge transmission chain from extension staff to model farmers to FREG members and other farmers to occur incomplete and therefore, lead to incomplete adoption. Consequently, in order to overcome this potential "failure to notice", farmers' attention needs to be drawn to each of the individual practices and the need for their integrated adoption. To do so, we designed a video intervention to complement the extension approach, which provides farmers with information on *why* each component is important, that is, explanations about the underlying principles and mechanisms of ISFM.

Building on these considerations, we derive a set of hypotheses for the context of ISFM knowledge diffusion and adoption in our experimental set-up. Firstly, we expect farmers to learn about ISFM through the extension intervention, and therefore hypothesize:

> H1: ISFM adoption and knowledge will increase through the extension activities, both of its individual components and the integrated package.

Further, we expect that farmers in treatment communities who are not actively involved in the extension activities benefit from information spillovers that occur via farmer-to-farmer communication or by observing neighbors' behavior, and therefore assume that:

> H1a: Due to information spillovers, ISFM adoption and knowledge will also increase among farmers not directly involved in extension activities.

Yet, farmers that "only" learn via informational spillovers are more likely to pick-up incomplete pieces of information (primarily what they consider most important, or what is easier to grasp), which lets us hypothesize:

> H1b: Since information spillovers occur incompletely, increases in ISFM adoption and knowledge will be lower for farmers not directly involved in extension activities.

We expect the additional video treatment to make farmers aware of potentially neglected knowledge dimensions, which is particularly beneficial for those who do not directly learn via extension. Thus, we hypothesize:

> H2: The additional video intervention counteracts incomplete information spillovers and therefore leads to higher ISFM knowledge and adoption.
>
> H2a: The additional "video effect" will be stronger for farmers that are not directly involved in extension activities.

Ultimately, since we expect that more complete knowledge fosters adoption, we hypothesize that:

> H3: Increases in ISFM adoption are (partly) channeled through gains in ISFM knowledge triggered by the interventions.

3. Experimental Design

This study builds on an RCT design with two treatment arms and a control group. We used microwatersheds (mws) as units of randomization, which are common implementation units for natural resource related interventions in Ethiopia. These are water catchment areas, i.e. natural hydrological entities defined by the topography of the land, typically consisting of around 250 to 300 households in one or several communities that share a common rainwater outlet.

3.1 Treatment Description

The core elements of the extension intervention introduced by the ISFM+ project are the following: In each treatment mws, three Farmer Research and Extension Groups (FREG) were formed, each consisting of 16 or 17 members who are "interested farmers working in collaboration with research and extension in the process of technology generation, verification and adoption" (ATA, 2014: 15), leading to a total of around 50 FREG members per mws. It is important to note that FREG farmers were selected in a non-random manner by extension agents and village heads, based on farmers' interest and social involvement. The FREGs conduct regular meetings to discuss on agricultural topics, typically once or twice per month. Each group is led by three model farmers that were appointed based on their reputation and farming skills in a participatory process with FREG members and extension staff during kick-off workshops. Model farmers are trained by public extension agents on ISFM techniques and provided with necessary inputs. Some of the model farmers or FREG members may be replaced from season to season, but this is not defined in a fixed way. The central activity of model farmers is the establishment and maintenance of demonstration plots. These are on-farm trials on which the package of ISFM practices is applied, next to plots that are managed according to traditional farming practices. Hence, differences in biomass production and yields resulting from common in contrast to ISFM practices should become clearly visible. In each mws, farmer field days are conducted twice per harvest cycle, at critical stages around mid-season and just before harvest. During these field days, model farmers share and discuss their experience with FREG members regarding the implementation of ISFM and its results, extension agents are present to complement information. Field day activities are mainly targeted to FREG members, although in some communities other farmers do also participate.

The ISFM+ extension treatment mainly aims at creating awareness and know-how about ISFM through a knowledge sharing process from development agents to model farmers, and from model farmers to FREG members. Through that entry point, information should diffuse to the broader population of farmers in the communities. Hence, this model heavily relies on peer-to-peer learning.

The video intervention has been designed to provide an additional stimulus to adoption by exposing farmers to information about the ISFM concept with all its individual components, in order to overcome potential knowledge gaps on key dimensions of the approach. The movie is composed of two parts: A narrative and documentary part which presents the example of a farmer couple who has already successfully implemented the quickwin technologies and visibly increased yields, serving as (potential) role models for treated farmers. These main characters explain their experience with implementation, emphasizing benefits and successes, but also critically discussing their initial reluctance and problems they have faced. In the narratives, particular emphasis was put on the fact that ISFM is a package approach and therefore, practices need to be combined on the same plot. The main characters were carefully chosen to be as far as possible representative of the target audience in terms of socioeconomic status and living environment. Given the cultural, linguistic as well as agro-ecological differences between Tigray, Amhara and Oromia, three different farmer couples were selected (i.e. one for each region). Previous research has underlined the importance of tailoring information to specific local conditions, as well as framing messages in a way that the audience can relate to them, which is best achieved by featuring credible role models from similar backgrounds (Jensen and Oster, 2009; Chong and La Ferrera, 2009; Chong et al., 2012; Bernard et al., 2014; 2015; 2016; BenYishay and Mobarak, 2018). Notwithstanding, all three versions strictly follow the same script in order to convey equal messages. Yet, it should be noted that, in contrast to the other two regions, the version for Tigray did not feature any information on liming, since soils in the intervention areas in this region do not suffer from acidity. In intervention areas in the other two regions, however, soil acidity presents one of the most severe threats, which makes the promotion of lime crucial. The second component of the film consists of animations that visualize processes taking place in the soil – such as hydrological cycles, the "work" of roots, soil organic matter, microorganisms and nutrients. Complex soil processes and the relationship between the individual ISFM components, soil fertility and improved yields are presented in a simplified way. Ultimately, farmers should gain a better understanding on why the integrated use of all techniques is important to improve soil fertility and productivity.

3.2 Sampling and Randomization Strategy

Since the participatory extension approach draws on the establishment of community-based farmer groups and demonstration sites, a cluster randomization approach has been applied, with mws as sampling unit. The decision to use microwatersheds was driven by the fact that the Sustainable Land Management Programme (SLMP) as umbrella project of the ISFM+ has equally been implemented at the mws level.

The full sampling frame consists of 161 mws located in 18 Woredas, equally distributed among the three regions Tigray, Amhara and Oromia. The list of target mws was compiled based on the criteria (i) benefiting from the SLMP and (ii) no/minimal previous exposure to soil fertility interventions. From this list, a sample of treatment mws was drawn randomly – stratified by region and Woreda – so that in in each Woreda four beneficiary mws were selected, resulting in a total of 72 treatment mws. Of these 72 treatment mws, half were assigned to the additional video intervention. Consequently, 36 mws received the extension treatment only (in the following referred to as T1), and another 36 mws received the extension treatment plus the additional video intervention (T2). In the second stage, in all treatment and control mws 15 households were randomly drawn from administrative lists, resulting in a total sample of 2,416 households. Thus, in treatment mws, the proportions of non-FREG and FREG farmers in the sample should on average represent their distribution in the population.

Figure 1 graphically depicts our full original sample, consisting of 2,416 households.

[Figure 1 about here]

3.3 Treatment Implementation

The ISFM+ project was launched in mid-2015. Yet, in the first months of operation, the project was still in the consolidation phase, i.e. conducting planning and kick-off workshops, while the implementation of the above described extension intervention on a broad scale started in the 2016 main cropping season. Since then, extension activities in T1 and T2 communities are on-going, regionally aligned with the course of the main harvest cycle (meher season).

The video screenings were conducted in T2 communities in early 2017, around six weeks prior to the start of the main growing season. Typically, the video was shown in public spaces such as farmer training centers, health posts or schools, and followed by group discussions that were facilitated by extension agents. In each microwatershed, the 15 household from our sample were invited by village heads a few days prior to the screenings orally and with written invitation cards. In the case of double-headed households we invited both spouses, otherwise only household heads.[2]

4. Data and Empirical Strategy

4.1 Data

In order to assess the interventions' impact, two rounds of survey data were collected. A baseline survey took place in early 2016, shortly after the launch of the ISFM+ project in mid-2015.[3] The endline survey took place in early 2018 among the same rural households. Data in both rounds were collected during tablet-based face-to-face interviews with the household head or spouse, using a structured questionnaire. Our attrition rate was remarkably low, since 2,382 (98.6%) of the 2,416 baseline sample households could be re-interviewed during endline, and we cannot detect any non-random patterns in this.

Both surveys covered modules on household sociodemographic characteristics, income and assets, food security level, social relationships, farming practices and agricultural production

---

[2] After the endline data collection, the video became freely available for extension staff to be used in T1 as well as control communities as well.

[3] The timing of the survey may raise concerns that baseline data might be influenced by first project activities, which we cannot completely rule out. Yet, as described above, structured extension activities in all treatment mws started in 2016.

data for the preceding cropping season, as well as exposure to agricultural extension. During endline, we assessed most information in the same way, yet had to adapt the mode of measurement for some variables. Besides, we added detailed questions on awareness for and participation in ISFM+ activities. In addition to the household-level questionnaire, in endline we included two individual-level modules administered to the household head as well as the spouse (in case the household was not single-headed), covering questions on the video content, psychological variables as well as a detailed knowledge exam. For the knowledge part, we first assessed farmers' awareness by asking them which ISFM components they actively remember, and in a second step, letting enumerators read through a list of practices and record which techniques respondents remember by name.[4] Subsequently, questions on their underlying principles and purpose ("principles knowledge") as well as their mode of implementation ("how-to knowledge") were posed. We combined different types of asking knowledge questions, including open questions, multiple choice tasks and statements on which respondents needed to decide between correct or incorrect (or a neutral "don't know" option) to minimize fatigue effects (for details of the knowledge exam, see Appendix B1). Enumerators were intensively trained and supervised during a ten-day training period. Questionnaire contents were carefully translated into the three local languages Amharic, Afan Oromo and Tigrigna and pretested in several rounds.

In addition to the farm household survey, we administered two community level questionnaires to key informants at the Woreda and mws levels, in order to collect data on infrastructure, extension exposure, rainfall and temperature, as well as other contextual characteristics.

4.2 Balance at Baseline

Table A1 (appendix A) depicts descriptive statistics on selected variables at baseline using data of the balanced panel of 2,382 households, which includes a test for covariate balancing between the three treatment groups to verify the success of our randomization process. Results show that households in T1, T2 and C are largely balanced on a series of sociodemographic and economic indicators. Yet, although we do not detect differences between households in T1 and T2, we do find some imbalances between treatment and control households with respect to agricultural production related variables. The use of inorganic fertilizer appears more widespread among T1 farmers, while slightly more T2 farmers appear to plant

---

[4] Inspired by Kondylis et al. (2015), we included a placebo practice ("seeding in circles") in this list to get a sense for possible response bias, which does not appear to threaten our results since yes-answers regarding this practice are close to zero.

main crops (teff, wheat, barley, maize and sorghum), and use improved seeds.[5] In addition, farmers in both T1 and T2 more frequently plant their crops in lines instead of broadcast them, and have more frequently attended agricultural training. These imbalances may present indications of early-intervention effects, and need to be taken into account in our outcome estimation framework. Further, we learn that ISFM practices are not necessarily new to farmers, since some were already used prior to the intervention, yet mostly to a relatively low extent.

4.3 Key Outcome Variables

Since our key concern is the increase in ISFM adoption, we measure how many ISFM practices farmers adopted in the preceding harvest cycle (2017), and therefore assess the *number of ISFM technologies adopted*, ranging from 0 to 5. We use this variable despite the fact that lime is not relevant in one of the regions (and thus, farmers in Tigray do effectively not reach a value of 5), but provide robustness checks verifying that implications do not change if we exclude lime and employ a 0 to 4 measure instead. Since the complementary use of practices is key to ISFM, our second main outcome is the *integrated adoption of the full ISFM package*. We assess this with a binary variable that measures whether a farmer has used all four quick-win practices in combination on at least one plot.[6] Here we exclude the use of lime, since adoption would otherwise always be zero in Tigray. Yet, to check sensitivity of our results with respect to this definition, we also use an alternative measure that equals 1 when at least four out of five practices (including lime) are adopted, and a regionally-specific measure that requires all five practices to be adopted in Amhara and Oromia, but only four in Tigray. Although effect sizes naturally vary with the choice of this measure, results remain qualitatively unaltered. Furthermore, we are interested in the adoption of the individual ISFM quickwin components, which are *compost*, *blended fertilizer*, *improved seeds*, *line seeding* and *lime*. For each technology, we define a dummy variable taking the value of 1 if the household has applied the respective practice in the 2017 main cropping season on any of its plots, for any crop type.

We are further interested in the effect of our treatments on ISFM knowledge as potential impact pathway to adoption. We construct an *overall ISFM knowledge* score based on questions on each of the individual ISFM components (but exclude the questions on lime). It rang-

---

[5] Although the baseline measure of improved seeds is likely to be noisy and should rather be considered as proxy.

[6] Due to data availability, we restrict these analyses to plots planted with main crops, i.e. wheat, maize, teff, barley and sorghum, which are also the main focus of the interventions.

es from 0 to 1, with 1 standing for full knowledge, i.e. having answered all questions correctly. Since the number of questions is not the same for all practices, we first calculate a knowledge score for each component individually and then combine it to an overall score, so that each ISFM dimension is included with the same weight in that indicator. Farmers who were not aware of a practice in the first place, were immediately given a value of zero in the respective follow-up questions. Since the aim of the video treatment was to increase farmers' knowledge on *why* ISFM is important (and not on *how* to implement it), we also construct two individual indicators for *principles* and a *how-to knowledge*, depending on whether a question was on the purpose (the "why") of a technology, or its mode of implementation (the "how"), which also range from 0 to 1. For the how-to score, we weigh knowledge on how to correctly produce compost double, since this is a more complicated process than the implementation of the other ISFM practices. For the principles score we weigh all ISFM components equally, and include one indicator for the general understanding of the necessity to integrate organic and inorganic soil inputs, that is whether the respondent agrees with the statement "The soil needs both organic and inorganic inputs to be healthy and fertile".

4.4 Identification Strategy

In order to assess the effect of our experimental interventions on ISFM adoption and knowledge, we estimate regressions of the following form:

$$Y_{i1} = \alpha + \beta\, T_i1 + \lambda\, T_i2 + \varepsilon_j + \epsilon_{ij} \tag{1}$$

where Y_{i1} denotes the respective outcome variable for household i, measured at endline. T_i1 is a dummy variable indicating whether farm household i lives in a community assigned to the extension intervention, and T_i2 indicates whether household i lives in an extension community that has additionally been randomly selected for the video screening. ε_j is the group level error term, while ϵ_{ij} represents the error term independently and identically distributed between individuals within groups. Standard errors are clustered at the mws level.

Although treatment indicators should be orthogonal to further explanatory variables due to randomization, we will re-estimate all models including additional covariates in order to increase precision of our estimates, which seems even more advisable considering the few initial imbalances in our sample:

$$Y_{i1} = \alpha + \beta\, T_i1 + \lambda\, T_i2 + \gamma\, X_{i0} + \varphi\, \mathrm{W_j} + \nu\, Y_{i0} + \varepsilon_j + \epsilon_{ij} \tag{2}$$

In these models, X_{i0} represents a vector of control variables related to farmer and household characteristics captured at baseline, while with W_j, we add further contextual variables measured at the community level, like infrastructure, rainfall and temperature. If available, we include the baseline level of outcome Y_{i0} to the equation in order to reduce its overall variance, since we assume some degree of path dependency on previously gained experience with a technology.[7] This treatment effect model is appropriate in our case, since for some outcomes, baseline and endline measures are not completely identical, or baseline data is not available at all (knowledge). In addition, this specification has been shown to be more powerful than the difference-in-difference estimator in the presence of relatively low autocorrelation, which can at least be stated for some of our outcome variables (McKenzie, 2012a; De Brauw et al., 2018).[8] Further, in each model we run a test of equality between β and λ, in order to evaluate the additional impact of the video intervention.

Our main parameters of interest β and λ give the intent-to-treat effect (ITT), which measures the average effect of living in a randomly assigned T1 or T2 community, irrespective of actual treatment participation. ITT estimates are of particular interest for policy makers, since in reality compliance is never expected to be perfect, and in our case not even intended. Nevertheless, they are likely to underestimate the "true" treatment effect on compliers, called treatment-on-the-treated (ToT) or local average treatment effect (LATE) (Duflo et al., 2007). Since the decision to participate is considered endogenous, a natural solution to tackle self-selection bias seems to use the random assignment as an instrument for actual participation. Yet, in order for this procedure to be valid, three assumptions need to hold (Angrist and Imbens, 1994; Duflo et al., 2007): Firstly, exposure to the treatment must be random – a condition which is satisfied given our randomized design. Secondly, exposure to the treatment is strongly correlated with actual participation. Finally, outcome variables must not directly be affected by the random assignment, but only through actual attendance. However, since we expect – and even aim for – that due to the decentralized nature of our extension treatment,

---

[7] Baseline data is available for adoption of compost, blended fertilizer and lime. Regarding blended fertilizer, we additionally control for ex-ante use of any inorganic fertilizer, since during time of baseline, blended fertilizer was largely unavailable (only 1.3% of sample households used it in baseline). Instead, farmers used the widely available DAP fertilizer (over 70% in baseline). In the two years between baseline and endline, supply-side structures changed in the way that more blended fertilizer factories were set up in Ethiopia and NPS/NPK fertilizer blends partly replaced other inorganic fertilizer types. Line seeding can only be proxied, since it was assessed on a more general level during baseline, asking farmers how they *usually* plant crops, but not at the plot level. Pre-intervention adoption of improved seeds is proxied using the source of seed as indicator, while in endline we assessed this variable in a more rigorous way. Knowledge variables were not measured in the baseline survey.

[8] Autocorrelation for outcomes: blended fertilizer: .0231; lime: .1159; improved seeds: .1607; line seeding: .4146; compost: .4785; no. of practices adopted: .4987.

adoption is affected via informational spillovers from actively participating individuals (i.e. FREG members) to their peers in communities, we do not see this third condition to be fulfilled in our setting. In fact, only around 50 farmers per mws are FREG members (which roughly corresponds to 15-20% of a mws's population) and thus, direct beneficiaries of the extension treatment.[9]

Therefore, we opt for a different approach that accounts for potential self-selection into FREGs, reverting to the basic idea of propensity score matching (Cameron and Trivedi, 2005). In a first step, we estimate a probit regression predicting FREG membership:

$$K_i = \alpha + \beta X_i + \epsilon_i \tag{3}$$

where X_i denotes a vector of farmer and household covariates assumed to influence farmer $i's$ participation K_i, and ϵ_i is the error term. Based on this estimation, a propensity score is calculated, with which we match each real FREG member in the treatment group with a hypothetical FREG member from the control group. We do so by using the simple one-to-one nearest-neighbor matching, i.e. we match each actual FREG member with the most similar counterfactual in the control group.[10] Subsequently, we re-estimate treatment effects on the core outcome variables within the two mutually exclusive subsamples: the *FREG sample*, consisting of actual and predicted FREG members (in treatment and control mws, respectively) and the *non-FREG sample*, consisting of actual non-FREG members in treatment and potential non-FREG members in control mws. By comparing the effects of the treatment only among real FREG farmers with those who would potentially be FREG farmers if they had the chance, we aim at ruling out a portion of unobserved heterogeneity that stems from the mere fact of being a "better" farmer rather than FREG membership itself.

In order to assess the importance of additional knowledge as potential impact pathways to adoption, we apply a causal mediation analysis, following Imai et al. (2011) and de Brauw et al. (2018). The aim is to estimate the average effect of our treatments T_i that is occurring through changes in knowledge as a mediating variable $M_i(T_i)$ that are triggered by the treatment. The causal mediation effect can be written as

---

[9] In fact, when using the instrumental variable approach to assess the ToT for FREG members only, we find implausibly large estimates for the aggregated ISFM adoption measures. They are likely to be inflated since the first stage of this estimation is not very powerful due to low compliance as per this definition (cp. Duflo et al. 2007).

[10] We consider this the cleanest approach and find that, compared to other matching algorithms like kernel matching or nearest-neighbor with replacement, this leads to the lowest bias and highest similarity along the chosen matching variables.

$$\delta_i(\mathrm{t}) \equiv Y_i\big(t, M_i(1)\big) - Y_i\big(t, M_i(0)\big) \tag{4}$$

in which $\mathrm{t} = 0, 1$ denotes the treatment status. By holding the treatment status otherwise constant at t and therefore eliminating all other causal mechanisms, $\delta_i(\mathrm{t})$ isolates the change in the outcome Y_i that stems from changing the mediator M_i from the control to the treatment condition.

The direct effect of the treatment $\zeta_i(\mathrm{t})$, that is the portion of the treatment effect not explained by the mediator, can be identified by changing the treatment status from 0 to 1, while fixing the effect of the mediator at t:

$$\zeta_i(\mathrm{t}) \equiv Y_i\big(1, M_i(t)\big) - Y_i(0, M_i(t)) \tag{5}$$

When averaging over all observations, the average causal mediation effect (ACME) is given by $\delta(t)$, and the average direct effect (ADE) is estimated by $\zeta(t)$, while the sum of the two $\delta(t) + \zeta(t)$ represents the total average treatment effect.

Two ignorability assumptions have to be made in order to estimate the ACME and the ADE (Imai et al., 2011). Firstly, treatment assignment is assumed to be independent of potential outcomes and mediators, conditional on baseline confounders. This exogeneity assumption holds due to randomization. Secondly, when controlling for actual treatment status and observed pretreatment characteristics, the mediating variable is statistically independent of potential outcomes.[11] This strong assumption is what Imai et al. (2010) call sequential ignorability and implies that no unobserved confounders exist that affect both our outcome and the mediator. Yet, in exchange for this assumption, we can then proceed to estimate the ACME and ADE by sequentially estimating the following equations:

$$M_{i1} = \alpha_2 + \beta_2 T_i 1 + \lambda_2 T_i 2 + \gamma_2 X_{i0} + \varphi_2 W_j + \nu_2 Y_{i0} + \varepsilon_{2i} \tag{6}$$

$$Y_{i1} = \alpha_3 + \beta_3 T_i 1 + \lambda_3 T_i 2 + \xi M_i + \gamma_3 X_{i0} + \varphi_3 W_j + \nu_3 Y_{i0} + \varepsilon_{3i} \tag{7}$$

The ACME of knowledge for T1 is given by $\widehat{\beta_2}\hat{\xi}$, where β_2 represents the effect of T1 on the mediator variable, and ξ the effect of the mediator on the outcome measure. Similarly, $\widehat{\lambda_2}\hat{\xi}$, gives the ACME of knowledge for T2.[12] Due to sequential ignorability, non-correlation be-

---

[11] For a formal description of these assumptions, refer to Imai et al. (2010) or Imai et al. (2011).

[12] Note that this formal description of causal mediation analyses assumes to fit linear regressions, in which both the outcome and the mediating variable are continuous measures. When the outcome is binary (as it is the case for the integrated adoption of the full ISFM package), the product of coefficients does not correspond to the ACME (Imai et al., 2010; Hicks and Tingley, 2011). Methods to correctly estimate mediation effects for binary outcomes and continuous measures have been developed and will be applied accordingly (Hicks and Tingley, 2011).

tween the error terms ε_{2i} and ε_{3i} is assumed ($\rho = 0$). Yet, since we can reasonably think of potential unobservable confounders that affect both knowledge and adoption (e.g. farmers' level of motivation or commitment) and would consequently bias our ACME estimates, we perform sensitivity tests in which we relax the assumption of $\rho = 0$ and re-estimate equations (6) and (7) for different hypothetical values of ρ.

5. Results

In this section, we first present a descriptive overview on farmers' participation in the interventions. Subsequently, we present and discuss ITT results on the effects of our interventions. Finally, we examine the contribution of gains in knowledge as potential impact pathways to adoption.

5.1 Treatment Participation

Among the two treatment groups, 82 farmers (8% of T1 and T2) were active model farmers in the 2017 cropping season, that is, they were responsible for the implementation and maintenance of an ISFM demonstration plot and provided with inputs from the project. In addition to model farmers, we find 120 further farmers (around 12% of T1 and T2) that are active FREG members, meaning they belong to a group and have participated in field day activities along the course of the preceding season. In addition, 77 (8%) of the treatment farmers who are no FREG members state to have participated in a field day in 2017, plus 39 (3%) of control group farmers. Regarding the visit of demonstration plots, 55 treatment farmers (6% of T1 and T2) not belonging to a FREG report to have visited a demonstration site on their own behalf, i.e. independently of a field day, in addition to 39 farmers in the control group (3% of C). Consequently, although to a low extent, we find indications of treatment spillovers both within and across groups, which also means our ITT estimates might suffer from a slight downward bias due to "contamination" of the control group.

Compliance in the video intervention was remarkably high, 499 (94%) of T2 households attended the screenings. Considering that in double-headed households we invited both spouses to the sessions, compliance at the individual level was 83%, equal to 804 participants.

5.2 ISFM Adoption Decision

Aggregated adoption measures

Table 1 shows the ITT effects of the two treatment arms on our first core outcome, the number of adopted ISFM quickwin technologies (0-5) obtained with three different regression specifications.[13] Since the dependent variable essentially is a count variable, we estimate a Poisson model. Yet, taking into account that it can also be perceived as either an underlying continuous or ordered process, we also estimate a linear as well as an ordered probit to underline the robustness of our findings.

[Table 1 about here]

The results of all models indicate positive and highly statistically significant effects of both treatments on the number of adopted practices, which are robust to the inclusion of baseline control variables. Furthermore, all models lead to larger point estimates for T2 compared to T1. From the linear model we see that farmers in T1 adopt additional .693 practices, while households in T2 communities on average adopt .852 more practices than those in the control group. When we include further baseline covariates, these coefficients drop to .450 and .536 respectively (columns (1) and (2)).[14] Columns (3) and (4) present the average marginal effects (AME) of the Poisson coefficients, which are well in line with the point estimates of the linear model and can equally be interpreted as additional practices adopted. Finally, results of the ordered specifications provide further evidence for positive and highly statistically significant effects of both the extension-only and the extension-plus-video treatment (columns (5) and (6)).[15] Yet, across all specifications, p-values of the test for equality of T1 and T2 (.421, .315, .421, .397, .398, .300) indicate that the average difference between the two treatment groups with regard to the number of adopted ISFM practices is not statistically significant.

Table 2 depicts the main effects and AME of being assigned to T1 and T2 on the integrated adoption of the full quickwin package, using a probit regression. As outlined earlier, we define integrated adoption as having adopted all four practices (compost, blended fertilizer,

---

[13] In the following, we always relay on the 0-5 measure when referring to the number of adopted practices, i.e. including lime. Yet, implications of our results do not change when using the 0-4 measure (excluding lime), see table A2 (Appendix A).

[14] In the following, we will always revert to the estimates of the second specification when interpreting our results, since conditioning on further control variables – and in particular the baseline value of the respective outcome (respectively its best available proxy) – should unarguably increase precision of the results.

[15] Since the coefficients of the ordered probit regressions cannot be interpreted in a straightforward way, we stick to interpreting results obtained from the OLS and Poisson models, relying on the ordered models as robustness checks. In addition, the assumption of parallel regressions underlying ordered probit models is violated which makes these results less reliable (Cameron and Trivedi, 2009).

improved seeds, line seeding) together on at least one (main crop) plot.[16] The estimated ITT effects are positive and statistically significant. The AME indicate that households in T1 are on average 7.6 percentage points more likely than control group households to adopt the full set of practices on the same plot, while the likelihood for farmers in T2 is 10.2 percentage points above the control group mean. However, again we do not detect a statistically significant difference between the effect sizes of T1 and T2 (p-values of equality tests .475 and .347, columns (1) and (2)).

[Table 2 about here]

In order to test whether the estimated treatment effects might be driven by the 82 model farmers in our sample that have been trained by development agents and provided with inputs, we re-estimate the ITT models on the two adoption outcomes excluding these 82 model farmers. We find that all treatment effects remain highly statistically significant while only decreasing slightly in their magnitude and can therefore conclude that the interventions affect farmers in treatment communities beyond model farmers (table A3 and A5, Appendix A).

Adoption of individual components

In order to shed light on which components are the main drivers of increased ISFM adoption, we subsequently examine the effects of the two treatment arms on the decision to adopt each of the five practices individually.

We assess households' decision to adopt each quickwin technology using binary probit models for each practice.[17, 18] Table 3 presents the AME of being assigned to the two treatments on the decision to adopt compost, blended fertilizer, improved seeds, line seeding and lime. Our primary estimates indicate that T1 exerts a positive and statistically significant effect on the decisions to adopt compost, improved seeds, line seeding and lime. In contrast, we find no robust significant effect of the extension-only intervention on the adoption of blended fertilizer. For T2, we detect positive and significant treatment effects on compost, line seeding and lime, while effects for blended fertilizer and improved seeds do not remain significant with the inclusion of additional explanatory variables.

---

[16] Yet, in table A4 (Appendix A) we show that using the two alternative specifications of this measure does not alter the interpretation of our results.

[17] When assessing several binary outcomes within one regression framework, a multivariate probit model may be favored over five individual binary models, since it is usually more efficient (Cappellari and Jenkins 2003). However, we find very similar estimates and standard errors with the mvprobit, and therefore opt for using the binary probit models, which allow for easier computation of AME and inclusion of covariates.

[18] Regarding the specifications with control variables, note again that baseline use of improved seeds and line seeding can only be approximated.

[Table 3 about here]

When assessing the effects of our interventions on five different, even if interrelated, outcomes, we are concerned that the observed effects in reality cannot be attributed to our interventions, but are rather detected by chance due to multiple outcome testing (Duflo et al. 2007). To account for the probability of false discoveries, we therefore follow Sankoh et al. (1997) and Aker et al. (2012) and use a version of the Bonferroni correction which takes into account inter-outcome correlations for families of outcomes (cp. Appendix B2). Although this procedure is less conservative than other corrections and presents a rather approximate fix, it nonetheless provides informative insights regarding the sensitivity of our findings (Sankoh, 1997; McKenzie, 2012b). With this form of adjustment, p-values of the estimated coefficients of both T1 and T2 increase above the .10 threshold for blended fertilizer and improved seeds, while results for compost, line seeding and lime remain significant for both treatment arms.

For these robust results, the estimated effect sizes of the extension-plus-video intervention arm are larger than those of the extension-only intervention, which is in line with the findings on the aggregated ISFM adoption measures. However, again we fail to detect any statistically significant differences between the effects of T1 and T2 on technology adoption (p-values of equality tests .260 and .202, columns (1) and (2); .472 and .872 columns (7) and (8); .416 and .267 columns (9) and (10)). For compost, the AME of T1 suggests that farmers in the extension treatment are on average 14.2 percentage points more likely to adopt than farmers in the control group. For T2, the AME indicate an increased likelihood of compost adoption of 18.3 percentage points (column (2)). Column (8) reports the AME for T1 and T2 regarding the adoption of line seeding, suggesting an increased likelihood to seed in lines of over 8 respectively 9 percentage points. Columns (9) and (10) show the AME for lime adoption, indicating that farmers in Amhara and Oromia who are assigned to T1 are on average around 20.6 percentage points more likely to adopt lime than those in the control group. Similarly, being assigned to T2 goes along with a likelihood to adopt lime that is about 23.8 percentage points above the control group mean. These effects seem substantial, considering that in the control group on average 4% adopt.

In summary, our results indicate significant ITT effects of the extension intervention on the adoption of ISFM, both on aggregated measures as well as on some of its individual components. Yet, despite consistently larger point estimates, we do not find significant evidence for an additional video effect.

Differential effects for FREG members

Until now, the ITT estimates represent the average effects of living in a treatment community. Considering the video treatment, we assume the ITT to be very close to the ToT effect, given the high compliance in the screenings (94% at household level). Yet, with regards to the extension intervention, the definition of actual compliance is not as straightforward. Recall that the core idea of the extension treatment is to spur ISFM adoption via peer-to-peer learning and the success of the intervention relies on information-sharing. For our analyses, this means that a large proportion of farmers in T1 and T2 have essentially not actively attended any extension activities. In fact, among the 1,069 farmers in T1 and T2, only 202 (19%) are FREG members, i.e. the primary target group of the extension activities, and of those 82 are model farmers. Consequently, the remaining 867 farmers (81%) only benefit from the extension treatment through spillover effects. Hence, we are interested in whether the extension treatment has an effect on ISFM adoption beyond FREG membership – or whether the estimated ITT is solely concentrated among FREG farmers –, and whether the additional video treatment might influence FREG members and non-FREG members differently. To do so, we use the PSM-inspired approach outlined in section 4.4.

Table 4 reports that in both subsamples, treatment effects of the two interventions regarding the number of adopted ISFM practices remain positive and highly significant.[19] Yet, both the linear and the Poisson specification indicate that the effects of the treatments are substantially larger in the FREG than in the non-FREG sample. While in the non-FREG sample, being assigned to T1 on average increases the number of applied technologies by .302, this coefficient is 1.193 in the FREG sample. Similarly, T2 is estimated to increase average adoption by .429 practices in the non-FREG sample, but by 1.054 technologies in the FREG sample (OLS results in columns (1) to (4)).[20]

Similarly, we examine the differential effects in the two subsamples regarding the integrated adoption of the technology package. Columns (9) to (12) of table 4 report that T1 does no longer carry a statistically significant coefficient in the non-FREG sample, while in the FREG sample, this effect stays significant at the 1% level, indicating that if T1 goes along with membership in a FREG, the likelihood of integrated adoption on average increases by around 26 percentage points in comparison to non-treated, yet potential FREG farmers in the

---

[19] We find a high level of common support with our matching algorithm, since only two treated observations are off support. See table A6 for first-stage PSM regression results and figure A1 for a histogram of the estimated propensity score (Appendix A).

[20] The AME estimates of the Poisson specification (columns (5) to (8) of table 4) are fairly close to the OLS estimates and are therefore not explicitly discussed.

control group. In contrast, if extension is complemented by the video intervention, the coefficient of the treatment variable (T2) stays highly statistically significant in both subsamples. In the FREG sample, extension-plus-video increases the likelihood of integrated adoption by 24 percentage points, a quantitatively very similar effect than that of extension-only. For non-FREG farmers, the likelihood to adopt all practices in combination on average appears over 7 percentage points higher compared to their hypothetical counterfactuals in non-treated communities.

[Table 4 about here]

We are confident to draw the following two conclusions: Firstly, the effect of the extension treatment seems substantially larger for FREG members – even after taking into account that they may be the better farmers anyways. Yet, it does still show a positive influence on non-FREG farmers when it comes to the number of adopted ISFM practices at the household level. However, most interestingly, our findings indicate that extension alone does not significantly affect non-FREG farmers when it comes to *integrated* adoption, i.e. comprehending the importance of combining the practices on the same plot. Yet, it seems that the video intervention compensates to some extent for the effect of the extension treatment on combined adoption of the practices, which is otherwise only absorbed via FREG membership.

5.3 ISFM Knowledge

Treatment effects on knowledge

Columns (1) and (2) of table 5 depict ITT estimates on the overall knowledge indicator. The positive and significant estimates in column (2) show that T1 on average seems to increase overall ISFM knowledge by almost 4 percentage points, while T2 increases farmers' knowledge by 7 percentage points in comparison to the control group mean. The p-value of .016 indicates that extension-plus-video exerts a significantly stronger effect on knowledge than extension alone and thus, points towards an additional effect of the video regarding ISFM knowledge formation. We also assess the ITT effects on the two distinct domains, principles- and how-to knowledge. Columns (3) and (4) of table 5 show that extension alone on average increases principles knowledge by more than 2 percentage points, whereas extension combined with video on average increases this knowledge indicator by almost 6 percentage points. The p-value of .016 indicates a significantly larger effect of T2 than of T1. How-to knowledge seems to be positively affected by both T1 and T2, with no statistical difference regarding their effect sizes (columns (5) and (6)).

[Table 5 about here]

Next we follow our earlier approach and disaggregate the sample into a FREG and a non-FREG sample (table 6). We find that the significant difference between the effect sizes of T2 and T1 on overall knowledge persists in the non-FREG, but not in the FREG sample, as the p-values of test of equality of T1 and T2 (.019, .018, and .432, .189) in columns (1) to (4) indicate. Regarding principles knowledge, extension alone does not show a significant effect in the non-FREG sample, while extension-plus-video does (columns (5) and (6)). In the FREG sample, T1 and T2 do both affect principles knowledge positively, yet the p-value of .089 (column (8)) points towards an additional effect of the video on principles knowledge formation even among this group of direct extension beneficiaries. Regarding knowledge on how to implement ISFM, both extension-only and extension-plus-video exert a positive influence for FREG members, with no statistical difference in their effect size (columns (11) and (12). Interestingly, we find that for non-FREG members the extension-plus-video treatment does increase how-to knowledge significantly stronger than extension-only (p-values of equality test .055 and .075, columns (9) to (10) of table 6). Further analyses reveal that this effect mainly stems from improved knowledge on how to produce compost among this group of farmers. This is fairly surprising, since the video did not convey any information on *how* to implement any of the practices. Yet, it may be that increased awareness and understanding of why ISFM is beneficial induced further knowledge-seeking processes on the mode of compost production among non-FREG farmers.

[Table 6 about here]

Causal mediation analysis

Our findings provide evidence that both extension-only as well as extension-plus-video increase farmers' knowledge about ISFM. Moreover, the video has triggered additional gains in knowledge, especially among those farmers that do not actively participate in the activities of the extension intervention. In particular, the video has shown success in understanding *why* the ISFM practices are important. Therefore, we seek to understand the contribution that these gains in ISFM knowledge make to the adoption decision. To do so, we conduct a causal mediation analysis as outlined in section 4.4, in which we use both the overall and the principles knowledge indicator as mediating variables.

Panel A of table 7 suggests positive and highly statistically significant effects of the two knowledge variables on both the number of adopted technologies as well as integrated adoption. Panel B presents the estimated ACME and ADE of T1 for both mediators and both adoption outcomes separately, Panel C the corresponding effects for T2. Regarding the number of adopted technologies, all ACME for the overall knowledge mediator are statistically significant and suggest that the portions of treatment effects explained by increases in overall knowledge are 11.9% for T1 and 17.9% for T2 (column (2) of Panels B and C). Gains in principles knowledge alone seem to account for 3.9% of the effect of T1 and for 7.7% of the effect of T2 (column (4) of Panels B and C). Looking at integrated ISFM adoption, the ACME of overall knowledge for both T1 and T2 are significant. Their effect sizes indicate that on average, an increase in overall ISFM knowledge triggered through T1 account for 13.3% of its total effect, while 18.0% of the effect of T2 seem to be transmitted through knowledge gains (column (6) of Panels B and C). Columns (7) and (8) of Panels B and C depict the ACME of principles knowledge regarding integrated adoption. For T1, the ACME is insignificant. Hence, though we find a small and marginally significant positive effect of T1 on principles knowledge (cp. table 5), these gains seem not to account for a portion of the effect of T1 on integrated ISFM adoption. In contrast, for T2, the ACME is significant, albeit small, and indicates that 7.0% of the effect of T2 on integrated adoption is driven by an increase in principles knowledge.

Since the sequential ignorability assumption we made to identify causal mediation effects is unjustifiably strong, we perform a sensitivity test to assess how severely our ACME estimates may be biased due to potential correlation of the error terms $\rho \neq 0$ of equations (6) and (7). Figures A2 to A9 in Appendix A depict the ACME for both mediators and both treatment variables as functions of varying values for ρ. Results show that only large negative correlations between the error terms would imply a strong impact of the knowledge mediators on both adoption outcomes. Yet, a positive correlation of error terms appears far more plausible, since unobservables determining additional unexplained knowledge should also positively affect unexplained adoption. In fact, when we estimate the correlation between error terms of equations (6) and (7) for both knowledge and both adoption variables, we find positive, but fairly small correlations never exceeding ρ =.003, so that our estimated ACMEs present upper bounds. This lets us conclude that, though gains in knowledge seem to explain a share of the treatment effects, their contribution appears rather modest, never explaining more than 18% of the estimated ITT.

6. Discussion and Conclusion

In this study we have assessed the impact of a decentralized extension model and an additional video intervention on knowledge and adoption of an ISFM technology package among smallholder farmers in Ethiopia using an RCT design. The success of decentralized approaches depends on information spillovers between farmers, since only a relatively small fraction of farmers actively takes part in extension activities. In line with selective attention theory, we argue that information loss occurs in the knowledge transmission process from actively participating farmers to their peers in communities. Since both information senders and receivers might not consider all aspects of ISFM important, they fail to teach and learn about it, especially when it comes to more complex aspects of a technology (package) that require more cognitive resources. Consequently, we expect ISFM adoption – both of its individual components, but in particular of the full integrated package – to be higher among active extension participants than among their non-participating peers in communities. To counteract information failures and draw farmers' attention to all component technologies as well as the necessity of applying them jointly, we designed a second intervention in form of a video explaining farmers the relevance of each individual practice and their synergistic potential.

We show that both extension-only and extension in combination with video lead to increased adoption of ISFM technologies – in particular of compost, line seeding and lime –, as well as increased integrated adoption, i.e. the use of the full set of practices together on the same plot. However, prima facie, we do not find evidence for a significant complementary effect of the video on adoption of the integrated package or any individual component, despite larger effect sizes of the combined over the extension-only treatment.

As expected, our findings reveal that treatment effects on the number of adopted technologies and adoption of the complete package are larger for those farmers who are members of a Farmer Research and Extension Group, the core component of the extension treatment. Yet, we still find positive and significant effects of T1 and T2 on non-FREG farmers in treated communities regarding the number of adopted practices. This points towards the existence of information spillovers from FREG farmers to their peers, that occur either through active information-sharing or through observation and imitation. Yet, since some farmers in treatment communities state to have attended a field day or visited a demonstration plot on their own behalf, even though they do not belong to a FREG, we rerun the analyses excluding all treatment farmers that have participated in any extension activity in any way and find that our positive treatment effects persist. These results provide support for the rationale of decentralized extension models and contradict previous research finding weak evidence for diffusion effects (Rola et al., 2002; Feder et al., 2004; Tripp et al., 2005; Kondylis et al., 2017).

Yet, when it comes to the *integrated* adoption of all technologies on the same plot, the extension-only treatment seems to do little for non-FREG farmers, while extension in combination with video does positively affect integrated adoption also among this group.

We hypothesized that increases in ISFM adoption are (partly) caused by gains in ISFM knowledge induced through our interventions. In fact, both interventions lead to higher knowledge about ISFM, both regarding its underlying principles and the way of implementing the practices. In line with our expectations, results suggest that the video provides a significant additional effect on overall knowledge, and in particular on understanding *why* ISFM is beneficial. However, for non-FREG farmers, we find that gains in principles knowledge are only triggered via the extension-plus-video intervention, but not through extension alone. This provides evidence that the video intervention indeed contributed to counterbalance incomplete information transmission by drawing farmers' attention to dimensions of the ISFM technology package they might not have noticed before or that are not transmitted via farmer-to-farmer extension at all.

In fact, we find evidence that possessing ISFM knowledge is positively associated with adoption. A formal pathway analysis reveals that higher knowledge on ISFM does partly account for the ITT effects of our interventions on ISFM adoption. These results suggest that a better understanding of ISFM as a package consisting of technologies that are all important and complement each other might indeed have positively influenced the decision to adopt the full package in an integrated manner, especially for those farmers that are excluded from the extension activities but did take part in the video intervention.

Yet, all in all we conclude that increases in knowledge only moderately explain the effects of our experimental treatments, which is in line with previous studies finding effects of extension on adoption that are only modestly explained by gains in knowledge (Kondylis et al., 2017; De Brauw et al., 2018). The limited explanatory effect of knowledge as impact channel might partly be attributed to imperfect measurement that knowledge assessments are frequently prone to, in the sense that with our questions we missed to capture some adoption-relevant dimensions of knowledge the treatments might have altered (Laajaj and Macours, 2017). Yet, as Kondylis et al. (2017) argue in the context of adoption of sustainable land management practices, knowledge constraints might simply not be the most decisive barrier to adoption, but rather a lack of awareness for their productivity benefits. In line with this, our treatments might have played a more crucial role in influencing farmers' awareness for the environmental and yield-enhancing benefits of ISFM, which has shown to be an important driver of adoption in the literature (Knowler and Bradshaw, 2007). Testimonies of the farmers about improvements of yields and their livelihoods presented in the video might have further increased the credibility of information.

Interestingly, for the group of non-FREG farmers, our results suggest that the additional video intervention triggered gains in knowledge on *how* to implement ISFM practices, albeit no explicit how-to messages were conveyed in the video. Further analyses reveal that these gains mostly stem from improved knowledge on the process of compost production, probably the most complex ISFM component. A possible explanation is that the video spurred how-to knowledge seeking processes. Increased awareness for ISFM and understanding why it is beneficial might have encouraged farmers to gather information on its mode of implementation, in particular on compost. This fits our argumentation in line with selective attention theory that additional information is especially needed for more complex technologies that farmers might otherwise disregard if they are not sufficiently convinced of their importance. In line with van Campenhout et al. (2017), another plausible explanation is that the video triggers affirmative processes, activating and making farmers feel more confident about latent knowledge they already possess, even in areas not explicitly mentioned in the video.

All in all, providing information via video seems a valuable method to complement decentralized extension activities in order to raise awareness for and knowledge about agricultural technologies. It might be especially beneficial for more marginalized farmers that are excluded from farmer groups and more likely to be bypassed by the information diffusion chain. This result is particularly policy-relevant given that video screenings are relatively simple to conduct, also in more remote geographical areas that typically benefit less from regular extension. The high compliance in the video screenings underline that farmers perceive video as an appealing format of information provision, which is in line with previous studies (e.g. Bernard et al., 2014). While in our case we only treated our 15 sample households in each community, the use of video could easily be scaled-up by conducting repeated screenings and admitting any interested farmer to participate. While most costs occur during video production and purchase of equipment, variable costs are low. Thus, video has the potential to achieve substantial outreach at a relatively low cost. However, its success in reaching those groups that are otherwise typically excluded still depends upon these farmers knowing about the screening to happen, which is certainly more difficult to achieve outside of an experimental setting in which we explicitly invited the sampled farmers. Screening videos during other community events or festivities would at least increase the chance of reaching more and different types of farmers.

Regarding the individual ISFM components, we find significant treatment effects on the adoption of compost, line seeding and lime, while impacts on adoption of blended fertilizer and improved seeds are not robust to conditioning on baseline covariates or adjusting p-values for multiple testing. When comparing adoption rates at baseline and endline (table A7, Appendix A), it becomes clear that particularly the use of the latter two technologies has substantially risen among all three groups, i.e. also among control farmers. This can be explained by the fact that farmers are probably less uncertain about these technologies. Blended fertilizer and improved seeds are relatively unambiguous practices, since the benefits of mineral fertilizer and quality seeds are more common knowledge than those of applying compost, line seeding or lime. Since we find evidence for information spillovers from active extension participants to non-participants within treatment communities, we can reasonably assume that information spillovers might also occur from farmers in treated to their peers in control communities, albeit to a lower extent; and that these spillovers happen more easily for less complex technologies.[21] In addition, the use of improved seeds and mineral fertilizer is promoted by the overall advisory system. In contrast, compost, line seeding and lime are less straightforward technologies, both in terms of their benefits and their application. The production of compost is not a trivial process and needs to be learned. The benefits of line seeding are often unclear to farmers. Since less seeds and fertilizer are used when crops are planted in rows, they commonly associate it with lower yields. In addition, distance recommendations for different crop types need to be learned. Lime is largely unknown by farmers, and for its (correct) use quantity and timing of application are crucial. In addition, the preparation of compost as well as seeding in lines are labor-intensive technologies, so that farmers need to be sufficiently convinced of their benefits in order to be willing to reallocate labor to these activities and gather information on how to implement them.

Consequently, it is advisable to concentrate information interventions primarily on more complex and less known practices, which seem to diffuse less easily between farmers.

---

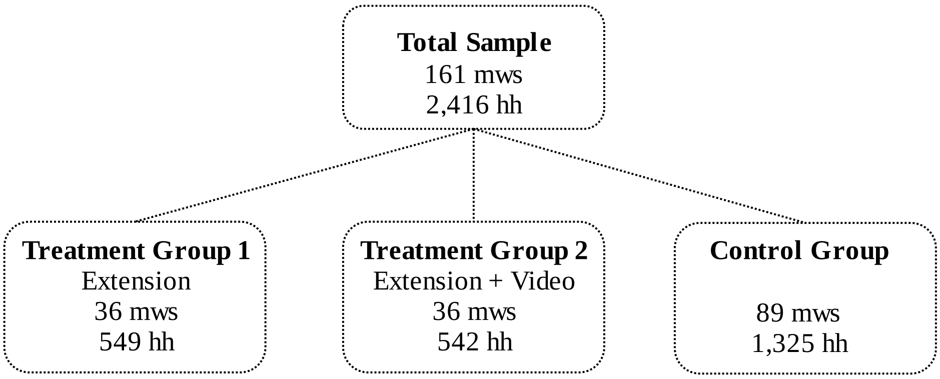
[21] In addition to communication spillovers from treated to non-treated farmers, we find that some, even if few, control farmers state to have participated in a field day or visited a demonstration plot in a treatment community (cp. section 5.1). Further, farmers in control mws might have also received ISFM information by communicating with extension staff at the Woreda or Kebele level. As pointed out earlier (5.1), potential spillovers to or contamination of the control group via any of these channels is likely to lead to an underestimation of our ITT effects.

A central question of experimental studies is to what extent findings are replicable or generalizable to other contexts or populations. For our case, we have to bear in mind that ISFM+ sites were chosen based on the precondition to already benefit from the Sustainable Land Management Programme. Building on SLMP achievements, it is thought that smallholders have higher capacities to address issues of soil fertility and productivity increase, which might not be the case in other contexts. Yet, although a replication of our interventions in other settings or with different populations may require adaptations, an advantage of our study regarding external validity is the comparatively large sample size spread over three regional states that in part differ quite substantially regarding agro-ecological, farming, climatic, cultural and other characteristics.

Our results add valuable insights for both researchers and policy-makers towards the effectiveness of decentralized, group-based extension approaches and the potential to close information gaps with complementary interventions, which can ultimately lead to more widespread adoption of system technologies such as ISFM. They appear particularly interesting given that the behavior of farmers to spread different types of inputs across plots rather than exploiting their complementary potential is largely understudied to date (Sheahan and Barrett, 2017).

Tables and Figures

Figure 1. Diagrammatic illustration of the final sample



Source: Own illustration.

Table 1. ITT effects on number of adopted ISFM technologies

	Number of ISFM technologies adopted					
	OLS		Poisson		Oprobit	
	(1)	(2)	(3)	(4)	(5)	(6)
T1	0.693***	0.450***	0.697***	0.464***	0.549***	0.537***
	(0.181)	(0.0755)	(0.175)	(0.0881)	(0.134)	(0.0791)
p-value	.000	.000	.000	.000	.000	.000
T2	0.852***	0.536***	0.834***	0.539***	0.682***	0.637***
	(0.175)	(0.0762)	(0.164)	(0.0848)	(0.134)	(0.0845)
p-value	.000	.000	.000	.000	.000	.000
Test T1=T2 (p-value)	.421	.315	.421	.397	.398	.300
Endline control mean	2.218					
Additional controls	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	0.073	0.536	.016	.122	.025	.227
Observations	2,382	2,382	2,382	2,382	2,382	2,382

Note: All models show treatment effects on number of practices adopted (0 to 5) measured at endline, controlling for baseline level of the outcome (respectively a proxy) in one specification each. Poisson results show average marginal effects (AME). Additional baseline control variables at household level are age, gender and education (in completed years) of household head; whether household had access to off-farm work or a non-farm family business; number of household members above age 14; walking distances to nearest farmer training center, paved road and market (in minutes); number of local organizations involved; access to irrigation, total land size in ha, tropical livestock units (TLU), a basic assets score, a food insecurity score, whether household is eligible for formal credit; whether household had a below-average preceding farming season; number of times household had contact with a development agent whether it has participated in agricultural training and grew main crops (teff, wheat, barley, maize, sorghum). Community level covariates are rainfall and temperature, whether microwatershed receives support from aid organizations other than GIZ, number of agricultural trainings offered in last season, presence of agricultural input dealer and seed enterprise in Kebele, distance to Woreda capital (in km) and two regional dummies for Oromia and Amhara. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table 2. ITT effects on integrated adoption of the full ISFM package

	Integrated adoption of full ISFM package	
	(1)	(2)
T1	0.104**	0.0757***
	(0.0424)	(0.0241)
p-value	.014	.002
T2	0.138***	0.102***
	(0.0430)	(0.0243)
p-value	.001	.000
Test T1=T2 (p-value)	.475	.347
Additional controls	No	Yes
Endline control mean	.152	
Pseudo R-squared	.023	.274
Observations	2,160	2,160

Note: Average marginal effects (AME) of probit models for outcomes measured at endline. Full quickwin package is a dummy variable taking 1 if all four practices (compost, blended fertilizer, line seeding and improved seeds) have been adopted together on at least one main crop plot. Additional control variables identical to those listed in notes of table 1. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table 3. ITT effects on adoption of individual ISFM components

	Adopted compost		Adopted blended fertilizer		Adopted improved seeds		Adopted line seeding		Adopted lime	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
T1	0.149***	0.142***	0.104*	0.0362	0.130**	0.0722**	0.161***	0.0872***	0.221***	0.206***
	(0.0560)	(0.0282)	(0.0605)	(0.0399)	(0.0579)	(0.0360)	(0.0553)	(0.0276)	(0.0417)	(0.0301)
Unadjusted p-values	.010	.000	.090	.368	.027	.045	.004	.002	.000	.000
Adjusted p-values	.000	.000	.263	.773	.109	.177	.014	.007	.000	.000
T2	0.221***	0.183***	0.115**	0.0375	0.129**	0.0580	0.209***	0.0927***	0.255***	0.238***
	(0.0541)	(0.0247)	(0.0547)	(0.0382)	(0.0584)	(0.0409)	(0.0570)	(0.0284)	(0.0415)	(0.0290)
Unadjusted p-values	.000	.000	.036	.328	.028	.159	.000	.002	.000	.000
Adjusted p-values	.000	.000	.112	.723	.113	.519	.000	.007	.000	.000
Robust to adjustment?	Yes		No		No		Yes		Yes	
Test T1=T2 (p-value)	.260	.202	.866	.976	.994	.749	.472	.872	.416	.267
Endline control mean	.405		.594		.573		.622		.040	
Pseudo R-squared	.027	.363	.010	.223	.014	.252	.035	.402	.134	.337
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2,382	2,382	2,382	2,382	2,382	2,382	2,382	2,382	1,464	1,464

Note: Average marginal effects (AME) of probit models for outcomes measured at endline, controlling for baseline level of the outcome (respectively a proxy) in one specification each. For lime, we exclude Tigray since it is not recommended in this region and adoption is zero. Additional control variables identical to those listed in notes of table 1. Bonferroni-adjusted p-values taking into account correlations between outcomes. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4. ITT effects on number of adopted ISFM technologies, FREG- and non-FREG samples separately

	Number of ISFM technologies adopted								Integrated adoption of full ISFM package			
	OLS				Poisson							
	Non-FREG sample		FREG sample		Non-FREG sample		FREG sample		Non-FREG sample		FREG sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
T1	0.573***	0.302***	1.166***	1.193***	0.577***	0.305***	1.169***	1.185***	0.0646	0.0276	0.263***	0.260***
	(0.189)	(0.0756)	(0.201)	(0.133)	(0.183)	(0.0882)	(0.194)	(0.141)	(0.0438)	(0.0208)	(0.0603)	(0.0466)
p-value	.003	.000	.000	.000	.002	.001	.000	.000	.140	.185	.000	.000
T2	0.755***	0.429***	1.070***	1.054***	0.734***	0.424***	1.086***	1.103***	0.110***	0.0736***	0.220***	0.244***
	(0.183)	(0.0778)	(0.187)	(0.141)	(0.170)	(0.0827)	(0.185)	(0.144)	(0.0417)	(0.0241)	(0.0687)	(0.0449)
p-value	.000	.000	.000	.000	.000	.000	.000	.000	.008	.002	.001	.000
Test T1=T2 (p-value)	.386	.135	.647	.291	.386	.171	.645	.487	.360	.081	.560	.765
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Endline control mean	2.106		2.284		2.106		2.284		.136		.232	
(Pseudo) R-squared	0.056	0.535	0.203	0.521	.013	.1281	.026	.067	.016	.296	.053	.293
Observations	1,984	1,984	398	398	1,984	1,984	398	398	1,772	1,772	388	387

Note: All models show outcomes measured at endline, controlling for baseline level of the outcome (respectively a proxy) in one specification each. Poisson and probit models (columns (5) to (12)) show average marginal effects (AME). Number of quickwin technologies adopted ranges from 0 to 5. Full quickwin package is a dummy variable taking 1 if all four practices (compost, blended fertilizer, line seeding and improved seeds) have been adopted together on at least one main crop plot. FREG- and non-FREG samples are constructed based on calculation of propensity scores predicting FREG membership. Additional control variables identical to those listed in notes of table 1. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table 5. ITT effects on different knowledge outcomes

	ISFM Knowledge					
	Overall		Principles knowledge		How-to knowledge	
	(1)	(2)	(3)	(4)	(5)	(6)
T1	0.0500***	0.0387***	0.0295*	0.0233*	0.0672***	0.0541***
	(0.0147)	(0.0128)	(0.0159)	(0.0139)	(0.0151)	(0.0130)
p-value	.001	.003	.066	.096	.000	.000
T2	0.0836***	0.0697***	0.0641***	0.0561***	0.0912***	0.0740***
	(0.0134)	(0.0114)	(0.0136)	(0.0115)	(0.0156)	(0.0133)
p-value	.000	.000	.000	.000	.000	.000
Test T1=T2 (p-value)	.020	.016	.029	.016	.168	.211
Endline control mean	0.448		0.522		0.382	
Additional controls	No	Yes	No	Yes	No	Yes
R-squared	0.034	0.207	0.012	0.147	0.048	0.209
Observations	2,334	2,334	2,334	2,334	2,334	2,334

Note: All models show treatment effects on household heads' knowledge scores measured at endline, using OLS regressions. Knowledge scores range from 0 to 1, are calculated based on the number of correct answers relative to the total number of questions in a respective domain. Additional control variables are age, gender, education (in completed years), whether respondent had access to off-farm work or a non-farm family business, whether household adopted the ISFM quickwin package at baseline, whether households has a cell phone and radio, number of times household had contact with a development agent, whether it has participated in agricultural training, number of local organizations involved, whether microwatershed receives support from aid organizations other than GIZ, and two regional dummies for Oromia and Amhara. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table 6. ITT effects on different knowledge outcomes, FREG- and non-FREG samples separately

	ISFM Knowledge											
	Overall				Principles knowledge				How-to knowledge			
	Non-FREG		FREG		Non-FREG		FREG		Non-FREG		FREG	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
T1	0.0329**	0.0232*	0.120***	0.111***	0.0171	0.0139	0.0774***	0.0601**	0.0427***	0.0307***	0.173***	0.173***
	(0.0146)	(0.0131)	(0.0236)	(0.0196)	(0.0169)	(0.0151)	(0.0281)	(0.0260)	(0.0137)	(0.0117)	(0.0264)	(0.0237)
p-value	.025	.079	.000	.000	.313	.360	.007	.022	.002	.009	.000	.000
T2	0.0646***	0.0532***	0.139***	0.139***	0.0459***	0.0423***	0.117***	0.110***	0.0711***	0.0554***	0.152***	0.158***
	(0.0129)	(0.0117)	(0.0193)	(0.0179)	(0.0135)	(0.0123)	(0.0260)	(0.0239)	(0.0149)	(0.0129)	(0.0184)	(0.0176)
p-value	.000	.000	.000	.000	.001	.001	.000	.000	.000	.000	.000	.000
Test T1=T2 (p-value)	.019	.018	.432	.189	.078	.054	.190	.089	.055	.075	.449	.459
Endline control mean	.434		.519		.510		.590		0.370		0.447	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.020	0.176	0.135	0.233	0.006	0.132	0.059	0.151	0.028	0.170	0.204	0.295
Observations	1,939	1,939	395	395	1,939	1,939	395	395	1,939	1,939	395	395

Note: All models show treatment effects on household heads' knowledge scores measured at endline, using OLS regressions. Knowledge scores range from 0 to 1, are calculated based on the number of correct answers relative to the total number of questions in a respective domain. FREG- and non-FREG samples are constructed based on calculation of propensity scores predicting FREG membership. Additional control variables identical to those listed in notes of table 5. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table 7. ITT and knowledge effects on number of adopted ISFM technologies and integrated adoption of the full ISFM package, ADE and ACME estimates of overall knowledge and principles knowledge as mediating variables

	Number of ISFM technologies adopted				Integrated adoption of full package			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Coefficient Estimates								
T1	0.559***	0.406***	0.656***	0.437***	0.085**	0.063**	0.099**	0.070***
	(0.160)	(0.0708)	(0.173)	(0.0750)	(0.042)	(.024)	(0.042)	(0.025)
T2	0.616***	0.456***	0.752***	0.504***	0.109**	0.086***	0.126***	0.095***
	(0.162)	(0.0729)	(0.171)	(0.0749)	(0.042)	(.024)	(0.043)	(0.024)
Overall knowledge score	2.903***	1.441***			0.397***	0.257***		
	(0.180)	(0.144)			(0.050)	(.056)		
Principles knowledge score			1.659***	0.764***			0.227***	0.119***
			(0.142)	(0.108)			(0.039)	(0.039)
(Pseudo) R-squared	0.215	0.564	0.146	0.550	.052	.285	.038	.276
Panel B: ACME and ADE Estimates of T1								
ACME (T1)	0.145***	0.055**	0.049*	0.018*	0.020**	0.010**	0.005	0.002
	(0.044)	(0.018)	(0.027)	(0.010)	(0.008)	(0.004)	(0.005)	(0.002)
ADE (T1)	0.566***	0.406**	0.663***	0.437***	0.095**	0.067**	0.110**	0.074**
	(0.151)	(0.073)	(0.163)	(0.077)	(0.045)	(0.027)	(0.045)	(0.028)
Total effect (T1)	0.712***	0.462***	0.712***	0.454***	0.111**	0.075**	0.112**	0.076**
	(0.161)	(0.074)	(0.166)	(0.079)	(0.045)	(0.028)	(0.045)	(0.028)
Share of treatment effect (T1) explained by knowledge	20.4%	**11.9%**	6.9%	**3.9%**	17.7%	**13.3%**	3.4%	2.9%
Panel C: ACME and ADE Estimates of T2								
ACME (T2)	0.243***	0.100***	0.106***	0.042***	0.038**	0.020***	0.017***	0.007**
	(0.042)	(0.018)	(0.025)	(0.010)	(0.008)	(0.005)	(0.005)	(0.003)
ADE (T2)	0.623***	0.456***	0.760**	0.504***	0.126**	0.094***	0.143**	0.103***
	(0.153)	(0.075)	(0.162)	(0.077)	(0.048)	(0.028)	(0.049)	(0.028)
Total effect (T2)	0.866***	0.556**	0.866**	0.547***	0.156***	0.111***	0.156***	0.109***
	(0.161)	(0.076)	(0.166)	(0.078)	(0.048)	(0.027)	(0.049)	(0.028)
Share of treatment effect (T2) explained by knowledge	28.0%	**17.9%**	12.3%	**7.7%**	24.8%	**18.0%**	11.1%	**7.0%**
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2,334	2,334	2,334	2,334	2,116	2,116	2,116	2,116

Note: Causal mediation estimates for two knowledge variables on number of practices adopted (OLS specification) and integrated adoption of full quickwin (probit specification, AME shown), measured at endline. Additional control variables identical to those listed in notes of table 1. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Appendix A

Table A1. Baseline descriptives and balance between treatment groups

	T1	T2	C	Overall	T1 vs. T2	T1 vs. C	T2 vs. C	N
Household Characteristics								
Age hh head	46.315	47.195	47.217	**47.008**	-0.880	-0.902	-0.022	2,382
	(0.666)	(0.766)	(0.528)	(0.371)	(0.999)	(0.844)	(0.924)	
Gender hh head *(0-female, 1-male)*	0.859	0.845	0.848	**0.850**	0.014	0.012	-0.003	2,382
	(0.017)	(0.017)	(0.012)	(0.009)	(0.024)	(0.021)	(0.021)	
No. of months hh head away from hh	0.100	0.098	0.117	**0.109**	0.002	-0.017	-0.018	2,382
	(0.036)	(0.028)	(0.021)	(0.015)	(0.046)	(0.041)	(0.035)	
HH head married *(dummy)*	0.833	0.832	0.830	**0.831**	0.002	0.003	0.002	2,382
	(0.020)	(0.019)	(0.012)	(0.009)	(0.028)	(0.024)	(0.023)	
Education hh head *(grades completed)*	2.198	2.407	2.034	**2.154**	-0.209	0.164	0.373	2,381
	(0.284)	(0.244)	(0.148)	(0.117)	(0.373)	(0.318)	(0.284)	
Literacy hh head *(dummy)*	0.535	0.560	0.573	**0.561**	-0.024	-0.038	-0.013	2,382
	(0.024)	(0.027)	(0.017)	(0.012)	(0.036)	(0.029)	(0.032)	
Main activity of hh head is farming *(dummy)*	0.894	0.887	0.888	**0.889**	0.008	0.006	-0.001	2,382
	(0.016)	(0.016)	(0.011)	(0.008)	(0.023)	(0.019)	(0.019)	
HH has non-farm family business *(dummy)*	0.202	0.180	0.180	**0.185**	0.022	0.022	0.000	2,382
	(0.025)	(0.027)	(0.016)	(0.012)	(0.037)	(0.029)	(0.031)	
Non-farm family business hh head *(dummy)*	0.104	0.098	0.107	**0.105**	0.005	-0.004	-0.009	2,381
	(0.015)	(0.018)	(0.012)	(0.008)	(0.024)	(0.019)	(0.021)	
HH participates in off-farm wage employment *(dummy)*	0.181	0.227	0.185	**0.194**	-0.045	-0.004	0.042	2,382
	(0.019)	(0.029)	(0.015)	(0.012)	(0.033)	(0.024)	(0.032)	
Off-farm wage employment hh head *(dummy)*	0.122	0.167	0.134	**0.139**	-0.044	-0.012	0.033	2,381
	(0.015)	(0.026)	(0.012)	(0.010)	(0.029)	(0.019)	(0.029)	
Household size	5.265	5.350	5.367	**5.340**	-0.085	-0.102	-0.017	2,382
	(0.134)	(0.129)	(0.081)	(0.061)	(0.183)	(0.155)	(0.151)	
No. of persons in hh above age 14	3.083	3.115	3.032	**3.062**	-0.032	0.051	0.083	2,382
	(0.086)	(0.068)	(0.042)	(0.034)	(0.109)	(0.095)	(0.079)	
No. of organizations involved	4.530	4.389	4.485	**4.474**	0.140	0.044	-0.096	2,382
	(0.155)	(0.141)	(0.111)	(0.077)	(0.208)	(0.189)	(0.178)	
Social capital score *(0-1)*	0.799	0.788	0.794	**0.794**	0.011	0.005	-0.006	2,382
	(0.012)	(0.010)	(0.007)	(0.005)	(0.016)	(0.014)	(0.012)	
Basic assets score *(0-4)*	1.798	1.907	1.823	**1.836**	-0.109	-0.025	0.084	2,382
	(0.070)	(0.077)	(0.048)	(0.035)	(0.104)	(0.083)	(0.089)	
Tropical livestock units (TLU)	3.257	3.481	3.415	**3.394**	-0.224	-0.158	0.066	2,382
	(0.211)	(0.197)	(0.149)	(0.105)	(0.283)	(0.257)	(0.245)	
TV owned *(dummy)*	0.011	0.019	0.017	**0.016**	-0.008	-0.006	0.002	2,382

	(0.005)	(0.009)	(0.004)	(0.003)	(0.010)	(0.006)	(0.010)	
Radio owned *(dummy)*	0.270	0.285	0.299	**0.289**	-0.015	-0.028	-0.013	2,382
	(0.022)	(0.025)	(0.018)	(0.012)	(0.033)	(0.028)	(0.030)	
Cellphone owned *(dummy)*	0.531	0.531	0.511	**0.520**	0.000	0.020	0.020	2,382
	(0.027)	(0.033)	(0.021)	(0.015)	(0.042)	(0.034)	(0.039)	
Contracted any credit *(dummy)*	0.383	0.350	0.320	**0.341**	0.034	0.063**	0.030	2,382
	(0.025)	(0.031)	(0.017)	(0.013)	(0.040)	(0.030)	(0.035)	
No. of hypothetical credit sources	1.711	1.717	1.747	**1.732**	-0.006	-0.036	-0.030	2,376
	(0.074)	(0.081)	(0.046)	(0.035)	(0.109)	(0.087)	(0.093)	
Eligible for formal credit *(dummy)*	0.711	0.728	0.746	**0.734**	-0.017	-0.035	-0.018	2,382
	(0.033)	(0.043)	(0.024)	(0.018)	(0.054)	(0.041)	(0.048)	
Received remittances *(dummy)*	0.089	0.098	0.108	**0.102**	-0.009	-0.019	-0.010	2,382
	(0.012)	(0.017)	(0.010)	(0.007)	(0.020)	(0.015)	(0.019)	
Received support from social programs *(dummy)*	0.163	0.212	0.219	**0.204**	-0.049	-0.056	-0.007	2,382
	(0.034)	(0.049)	(0.030)	(0.021)	(0.059)	(0.045)	(0.057)	
Food insecurity score *(0-food secure, 12-severly food insecure)*	2.259	2.076	2.471	**2.335**	0.184	-0.211	-0.395	2,382
	(0.221)	(0.221)	(0.149)	(0.108)	(0.310)	(0.264)	(0.264)	
Agricultural Production Characteristics								
Total land size in ha	1.443	1.493	1.462	**1.465**	-0.050	-0.019	0.031	2,317
	(0.127)	(0.123)	(0.103)	(0.069)	(0.173)	(0.162)	(0.160)	
Grows main crop *(dummy)*	0.942	0.953	0.930	**0.938**	-0.011	0.013	0.023*	2,382
	(0.011)	(0.009)	(0.011)	(0.007)	(0.014)	(0.015)	(0.014)	
Irrigation used *(dummy)*	0.169	0.195	0.196	**0.189**	-0.026	-0.027	-0.001	2,382
	(0.035)	(0.035)	(0.022)	(0.016)	(0.049)	(0.041)	(0.041)	
Compost used *(dummy)*	0.337	0.386	0.366	**0.364**	-0.049	-0.029	0.020	2,382
	(0.044)	(0.042)	(0.032)	(0.022)	(0.061)	(0.054)	(0.053)	
Blended fertilizer applied *(dummy)*	0.009	0.021	0.014	**0.014**	-0.012	-0.004	0.007	2,382
	(0.005)	(0.010)	(0.005)	(0.004)	(0.011)	(0.007)	(0.011)	
Improved seeds used (proxy) *(dummy)*	0.146	0.163	0.110	**0.130**	-0.016	0.036	0.052**	2,382
	(0.021)	(0.020)	(0.012)	(0.009)	(0.029)	(0.024)	(0.023)	
Plants crops usually in lines *(dummy)*	0.524	0.526	0.388	**0.450**	-0.001	0.136**	0.137*	2,382
	(0.059)	(0.062)	(0.036)	(0.028)	(0.085)	(0.068)	(0.071)	
Lime applied *(dummy)*	0.009	0.009	0.007	**0.008**	0.000	0.002	0.003	2,382
	(0.005)	(0.005)	(0.002)	(0.002)	(0.007)	(0.005)	(0.005)	
Reduced tillage practiced *(dummy)*	0.091	0.051	0.064	**0.067**	0.040	0.027	-0.013	2,382
	(0.029)	(0.015)	(0.012)	(0.010)	(0.033)	(0.031)	(0.020)	
Manure applied *(dummy)*	0.506	0.478	0.482	**0.487**	0.027	0.023	-0.004	2,382
	(0.029)	(0.030)	(0.023)	(0.016)	(0.042)	(0.037)	(0.038)	
DAP applied *(dummy)*	0.763	0.733	0.663	**0.701**	0.030	0.100**	0.071	2,382
	(0.034)	(0.039)	(0.033)	(0.022)	(0.052)	(0.047)	(0.051)	
Urea applied *(dummy)*	0.709	0.662	0.601	**0.639**	0.048	0.108**	0.061	2,382
	(0.043)	(0.041)	(0.032)	(0.023)	(0.059)	(0.053)	(0.052)	
Intercropping practiced	0.187	0.174	0.159	**0.169**	0.013	0.028	0.015	2,382

(dummy)								
	(0.034)	(0.032)	(0.022)	(0.016)	(0.046)	(0.040)	(0.038)	
Multi-purpose trees planted *(dummy)*	0.143	0.144	0.127	**0.134**	-0.001	0.015	0.016	2,382
	(0.024)	(0.020)	(0.014)	(0.010)	(0.031)	(0.027)	(0.024)	
Green manure crops grown *(dummy)*	0.030	0.026	0.018	**0.022**	0.003	0.012	0.009	2,382
	(0.009)	(0.011)	(0.005)	(0.004)	(0.014)	(0.010)	(0.011)	
Experienced shock in last season *(dummy)*	0.802	0.794	0.834	**0.818**	0.008	-0.032	-0.040	2,382
	(0.033)	(0.030)	(0.019)	(0.014)	(0.044)	(0.037)	(0.035)	
Below average preceding farming season *(dummy)*	0.459	0.448	0.506	**0.482**	0.011	-0.046	-0.058	2,382
	(0.058)	(0.054)	(0.038)	(0.027)	(0.078)	(0.068)	(0.066)	
Av. perception of parcel fertility *(1-below av., 2-average, 3-above av.)*	1.925	1.910	1.921	**1.919**	0.015	0.004	-0.011	2,381
	(0.042)	(0.035)	(0.028)	(0.020)	(0.054)	(0.050)	(0.044)	
Av. perception of change in parcel fertility over last 3 years *(1-decreased, 2-same, 3-increased)*	1.877	1.902	1.875	**1.881**	-0.025	0.001	0.026	2,381
	(0.071)	(0.064)	(0.038)	(0.030)	(0.094)	(0.080)	(0.074)	
No. of times talked to DA in the last 12 months	5.757	6.452	5.056	**5.525**	-0.694	0.702	1.396	2,380
	(0.776)	(0.963)	(0.463)	(0.377)	(1.228)	(0.897)	(1.059)	
Attended agric. training in the last 12 months *(dummy)*	0.304	0.336	0.227	**0.269**	-0.033	0.077**	0.110***	2,382
	(0.032)	(0.032)	(0.019)	(0.015)	(0.045)	(0.037)	(0.037)	
Observations	540	529	1,313	2,382	1,069	1,853	1,842	

Note: Main crops are teff, wheat, barley, maize, sorghum. Basic asset score comprises the following characteristics: household has modern roof, improved stove, modern lighting, toilet facility. Conversion factors used for calculation of TLU: camel=1, horse=0.8, oxen/cow/mule=0.7, donkey=0.5, goat/sheep=0.1, chicken=0.01. Model farmers included. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table A2. ITT effects on number of adopted ISFM technologies, using alternative 0-4 measure

	Number of ISFM quickwin technologies adopted (0-4)					
	OLS		Poisson		Oprobit	
	(1)	(2)	(3)	(4)	(5)	(6)
T1	0.551***	0.328***	0.553***	0.338***	0.454***	0.404***
	(0.167)	(0.0761)	(0.162)	(0.0881)	(0.137)	(0.0889)
p-value	.001	.000	.001	.000	.001	.000
T2	0.673***	0.388***	0.660***	0.393***	0.568***	0.482***
	(0.158)	(0.0761)	(0.150)	(0.0864)	(0.135)	(0.0914)
p-value	.000	.000	.000	.000	.000	.000
Test T1=T2 (p-value)	.476	.475	.476	.540	.453	.444
Endline control mean	2.195					
Additional controls	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	0.053	0.508	.011	.107	.019	.229
Observations	2,382	2,382	2,382	2,382	2,382	2,382

Note: All models show treatment effects on number of practices adopted (0 to 4, excluding lime) measured at endline, controlling for baseline level of the outcome (respectively a proxy) in one specification each. Poisson results show average marginal effects (AME). Additional control variables identical to those listed in notes of table 1. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table A3. ITT effects on number of adopted ISFM technologies, excluding model farmers

	Number of ISFM technologies adopted					
	OLS		Poisson		Oprobit	
	(1)	(2)	(3)	(4)	(5)	(6)
T1	0.593***	0.394***	0.593***	0.400***	0.467***	0.473***
	(0.180)	(0.0728)	(0.173)	(0.0851)	(0.134)	(0.0773)
p-value	.001	.000	.000	.000	.001	.000
T2	0.749***	0.491***	0.729***	0.493***	0.607***	0.595***
	(0.179)	(0.0790)	(0.166)	(0.0841)	(0.139)	(0.0896)
p-value	.000	.000	.001	.000	.000	.000
Test T1=T2 (p-value)	.436	.259	.432	.287	.380	.215
Endline control mean	2.218					
Additional controls	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.056	.531	.012	.122	.020	.225
Observations	2,300	2,300	2,300	2,300	2,300	2,300

Note: All models show treatment effects on number of practices adopted (0 to 5) measured at endline, controlling for baseline level of the outcome (respectively a proxy) in one specification each. Poisson results show average marginal effects (AME). Additional baseline control variables at household level are age, gender and education (in completed years) of household head; whether household had access to off-farm work or a non-farm family business; number of household members above age 14; walking distances to nearest farmer training center, paved road and market (in minutes); number of local organizations involved; access to irrigation, total land size in ha, tropical livestock units (TLU), a basic assets score, a food insecurity score, whether household is eligible for formal credit; whether household had a below-average preceding farming season; number of times household had contact with a development agent whether it has participated in agricultural training and grew main crops (teff, wheat, barley, maize, sorghum). Community level covariates are rainfall and temperature, whether microwatershed receives support from aid organizations other than GIZ, number of agricultural trainings offered in last season, presence of agricultural input dealer and seed enterprise in Kebele, distance to Woreda capital (in km) and two regional dummies for Oromia and Amhara. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table A4. ITT effects on integrated adoption of the full ISFM package, using alternative measures

	Integrated adoption of full ISFM package (at least 4 out of 5)		Integrated adoption of full ISFM package (region-specific)	
	(1)	(2)	(3)	(4)
T1	0.153***	0.114***	0.0626***	0.0600***
	(0.0434)	(0.0266)	(0.0170)	(0.0132)
p-value	.000	.000	.000	.000
T2	0.189***	0.143***	0.0757***	0.0602***
	(0.0422)	(0.0248)	(0.0205)	(0.0147)
p-value	.000	.000	.000	.000
Test T1=T2 (p-value)	.455	.334	.469	.972
Additional controls	No	Yes	No	Yes
Endline control mean	.157		.033	
Pseudo R-squared	.040	.278	.043	.251
Observations	2,160	2,160	2,160	2,160

Note: Average marginal effects (AME) of probit models for outcomes measured at endline, controlling for baseline level of the outcome (respectively a proxy) in one specification each. In columns (1) to (4), full quickwin package is a dummy variable defined as adopting at least four out of five practices (including lime). In columns (5) to (8), full package is a dummy variable defined as adopting all five practices in Amhara and Oromia (including lime), but only four in Tigray, since lime is not relevant there. Additional control variables identical to those listed in notes of table 1. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table A5. ITT effects on integrated adoption of the full ISFM package, excluding model farmers

	Integrated adoption of full ISFM package	
	(1)	(2)
T1	0.078*	0.057**
	(0.0422)	(0.0228)
p-value	.064	.012
T2	0.108**	0.083***
	(0.0435)	(0.0257)
p-value	.013	.001
Test T1=T2 (p-value)	.542	.354
Additional controls	No	Yes
Endline control mean	.152	
(Pseudo) R-squared	.015	.274
Observations	2,078	2,078

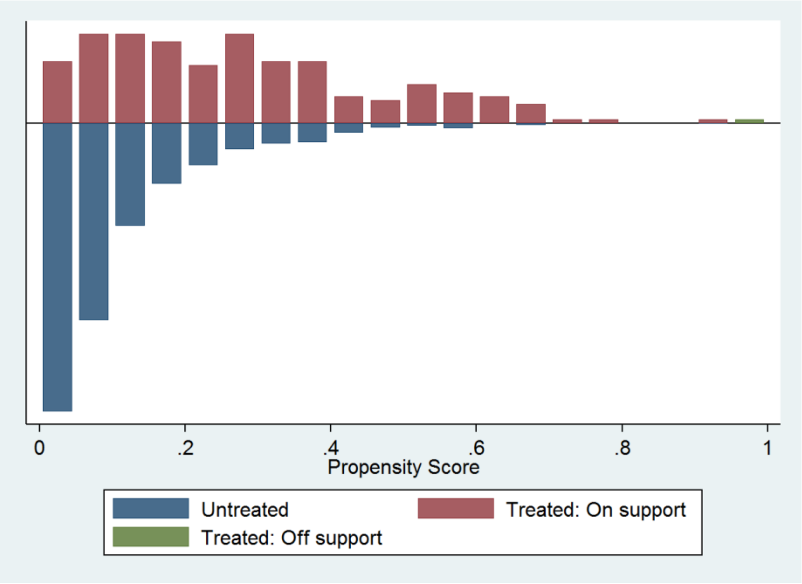
Note: Average marginal effects (AME) of probit models for outcomes measured at endline. Full quickwin package is a dummy variable taking 1 if all four practices (compost, blended fertilizer, line seeding and improved seeds) have been adopted together on at least one main crop plot. Additional control variables identical to those listed in notes of table 1. Tests of equality of T1 and T2 are Wald tests. Robust standard errors in parentheses, clustered at the microwatershed level. Significance levels indicated by asterisks as following: *** p<0.01, ** p<0.05, * p<0.1.

Table A6. Probit regression results for calculation of propensity score to predict FREG membership, used for matching with potential FREG members in control group

	FREG membership
Gender hh head *(0-female, 1-male)*	-0.518**
	(0.231)
Age hh head	-0.00378
	(0.00394)
No. of months hh head away from hh	-0.0490
	(0.0582)
HH head married *(dummy)*	0.322
	(0.226)
Education hh head *(grades completed)*	0.0369***
	(0.0137)
Non-farm family business hh head *(dummy)*	-0.224
	(0.144)
Off-farm wage employment hh head *(dummy)*	0.0902
	(0.139)
Houseold size	0.0142
	(0.0266)
Social capital score *(0-1)*	-0.0261
	(0.216)
Help-trust score *(0-1)*	-0.434*
	(0.246)
Father of hh head important in Kebele *(dummy)*	0.140
	(0.103)
Walking distance to nearest FTC *(in minutes)*	-0.000271
	(0.00178)
No. of times talked to DA in the last 12 months	0.0127***
	(0.00437)
Attended agric. training in the last 12 months *(dummy)*	0.783***
	(0.0966)
Basic assets score *(0-4)*	0.0289
	(0.0551)
Radio owned *(dummy)*	-0.190*
	(0.107)
Food insecurity score *(0-food secure, 12-severly food insecure)*	-0.0738***
	(0.0187)
Received support from social programs *(dummy)*	0.166
	(0.117)
Total land size in ha	0.152***
	(0.0398)
Tropical livestock unit (TLU)	-0.0262
	(0.0193)
No. of quickwin technologies adopted	0.184***
	(0.0481)
Grows main crop *(dummy)*	0.300
	(0.313)
Constant	-1.669***
	(0.479)
Pseudo R-squared	.186
Observations	1,486

Note: Probit results for calculation of propensity score to predict FREG membership in treatment groups. Robust standard errors in parentheses. Significance levels indicated by asteriks as following: *** p<0.01, ** p<0.05, * p<0.1.

Figure A1. Histogram of estimated propensity score, using nearest-neighbor matching



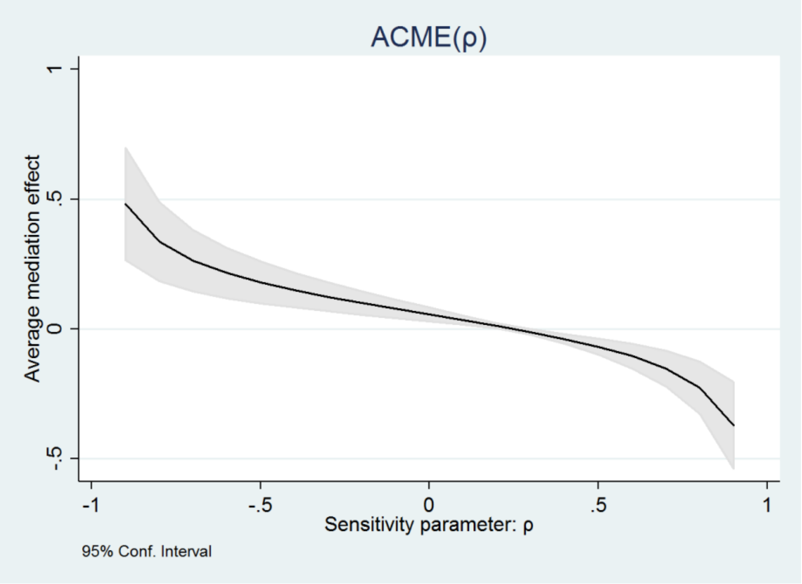


Figure A2. Sensitivity test ACME of overall knowledge (T1), no. of adopted practices

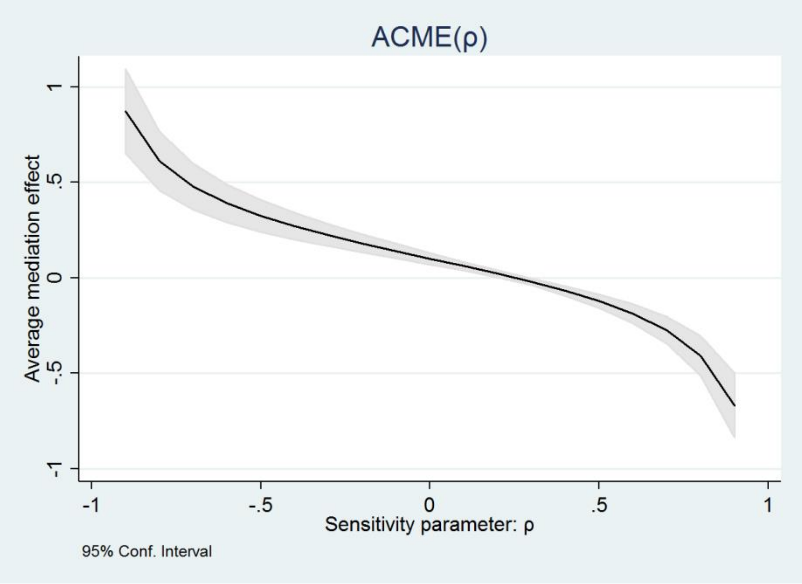


Figure A3. Sensitivity test ACME of overall knowledge (T2), no. of adopted practices

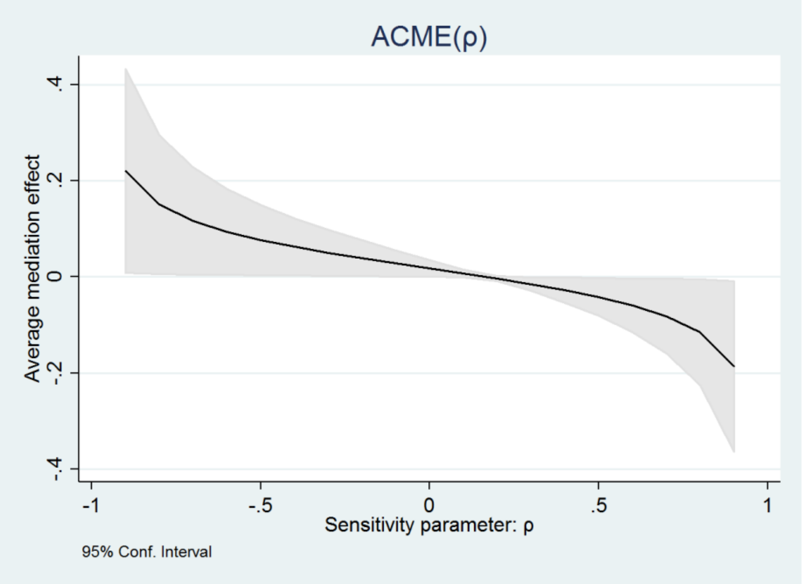


Figure A4. Sensitivity test ACME of prin. knowledge (T1), no. of adopted practices

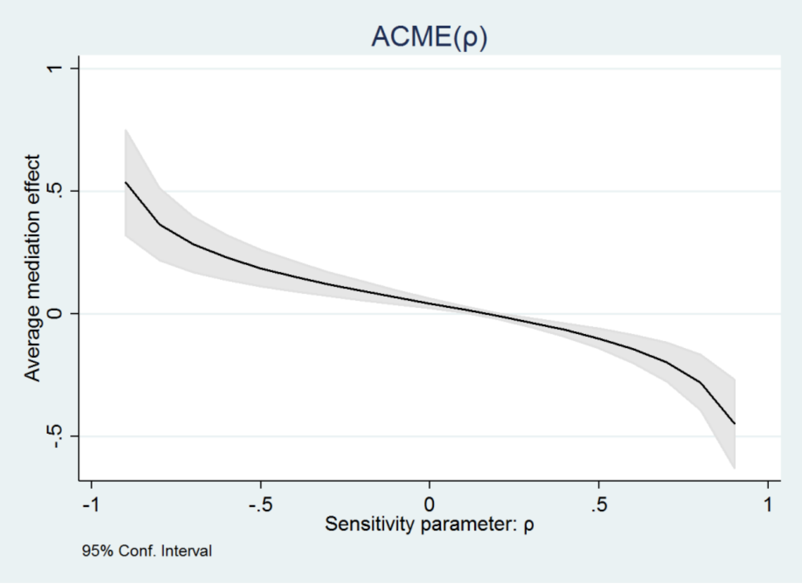


Figure A5. Sensitivity test ACME of prin. knowledge (T2), no. of adopted practices

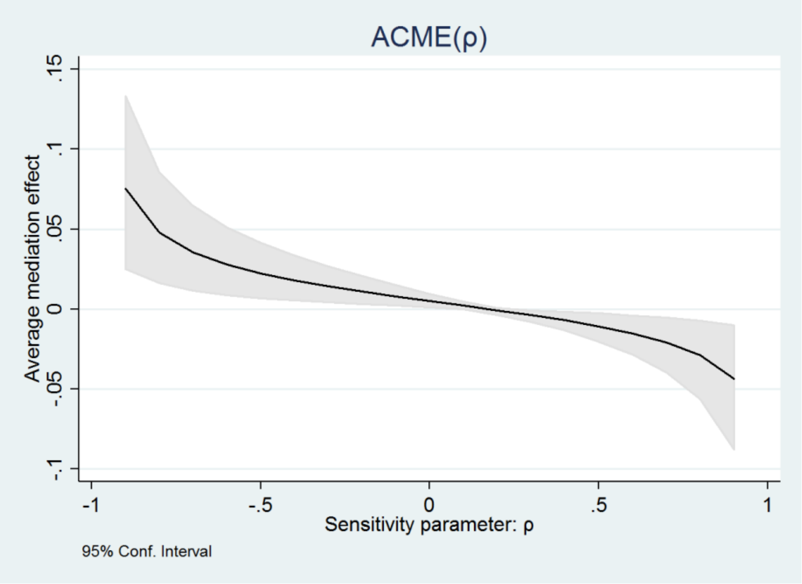


Figure A6. Sensitivity test ACME of overall knowledge (T1), integrated adoption

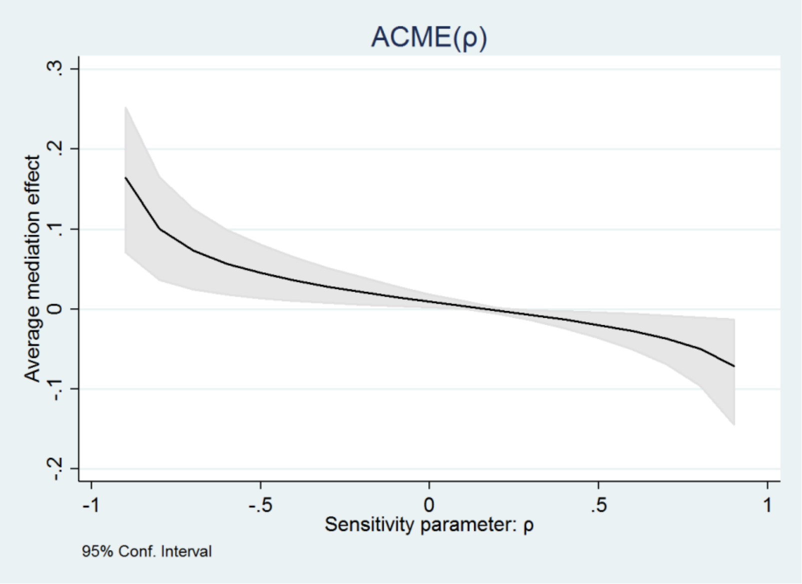


Figure A7. Sensitivity test ACME of overall knowledge (T2), integrated adoption

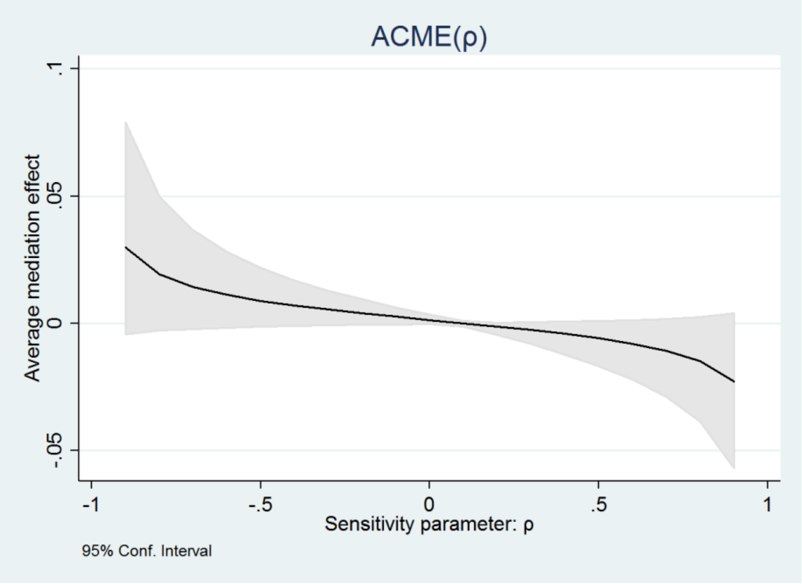


Figure A8. Sensitivity test ACME of princ. knowledge (T1), integrated adoption

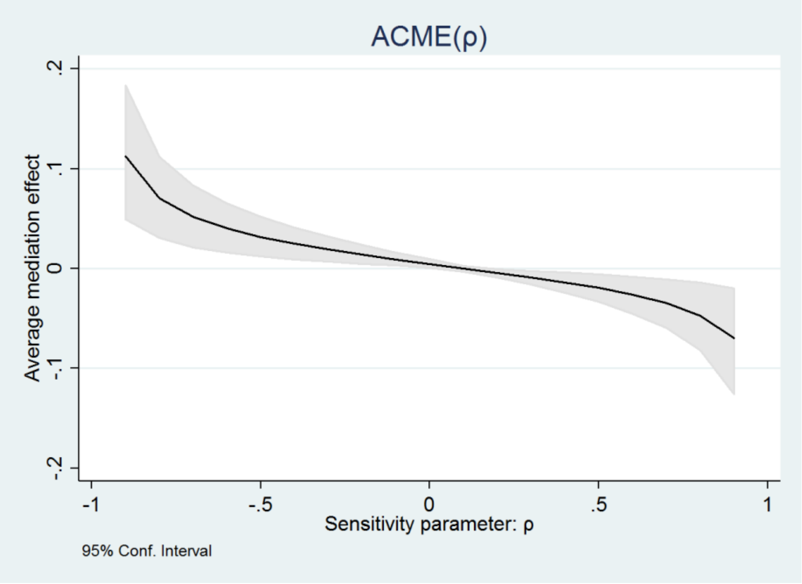


Figure A9. Sensitivity test ACME of princ. knowledge (T2), integrated adoption

Table A7. Average adoption of ISFM practices at baseline and endline, with differences between treatment groups

	Average baseline and endline outcomes by treatment group							
	T1	T2	C	Overall	T1 vs. T2	T1 vs. C	T2 vs. C	N
Adopted compost								
Baseline	0.337	0.386	0.366	0.364	-0.049	-0.029	0.020	2,382
	(0.044)	(0.042)	(0.032)	(0.022)	(0.061)	(0.054)	(0.053)	
Endline	0.557	0.629	0.405	0.490	-0.072	0.152**	0.224***	2,382
	(0.047)	(0.044)	(0.035)	(0.025)	(0.064)	(0.058)	(0.056)	
Adopted blended fertilizer								
Baseline	0.009	0.021	0.014	0.014	-0.012	-0.004	0.007	2,382
	(0.005)	(0.010)	(0.005)	(0.004)	(0.011)	(0.007)	(0.011)	
Endline	0.698	0.709	0.594	0.643	-0.011	0.104*	0.115**	2,382
	(0.049)	(0.041)	(0.035)	(0.024)	(0.064)	(0.060)	(0.054)	
Adopted improved seeds								
Baseline (proxy)	0.146	0.163	0.110	0.130	-0.016	0.036	0.052**	2,382
	(0.021)	(0.020)	(0.012)	(0.009)	(0.029)	(0.024)	(0.023)	
Endline	0.704	0.703	0.573	0.632	0.000	0.130**	0.130**	2,382
	(0.045)	(0.045)	(0.036)	(0.025)	(0.063)	(0.057)	(0.058)	
Adopted line seeding								
Baseline (proxy)	0.524	0.526	0.388	0.450	-0.001	0.136**	0.137*	2,382
	(0.059)	(0.062)	(0.036)	(0.028)	(0.085)	(0.068)	(0.071)	
Endline	0.787	0.826	0.622	0.705	-0.039	0.165***	0.204***	2,382
	(0.040)	(0.038)	(0.038)	(0.025)	(0.054)	(0.055)	(0.053)	
Adopted lime								
Baseline	0.014	0.011	0.012	0.012	0.003	0.002	-0.001	1,464
	(0.007)	(0.007)	(0.004)	(0.003)	(0.010)	(0.008)	(0.008)	
Endline	0.248	0.303	0.040	0.154	-0.055	0.208***	0.263***	1,464
	(0.046)	(0.051)	(0.011)	(0.021)	(0.068)	(0.047)	(0.052)	
Observations	540	529	1313	2382	1069	1853	1842	

Note: Since line seeding and improved seeds were assessed slightly differently during baseline, baseline values present proxies of the endline outcome. For lime, we exclude households in Tigray since it is not recommended in this region and adoption is zero. Robust standard errors in parentheses, clustered on the microwatershed level; significance levels indicated by asteriks as following: *** p<0.01, ** p<0.05, * p<0.1.

Appendix B

B1. Knowledge Exam

Known by Memory

K1.) What are the most important components of integrated soil fertility management?
(open question)

Known by Name

K2.) Which of the following technologies have you heard of before this interview?
(list of several ISFM technologies read out)

How-to Knowledge

K3.) Imagine you buy improved seeds for wheat. For how many cropping seasons could you reuse them until you should purchase new ones?

Up to four cropping seasons *(correct)*
Five to eight cropping seasons
I can use them endlessly, no need to purchase again
Don't know

K4.) What are the three most important ingredients if you want to produce good-quality compost? *(open question; correct if mentions at least one nitrogen- and one carbon-rich material)*

K5.) What is the optimal sequence of layers to produce improved compost?
(choose the correct out of three pictures)

K6.) In order to produce good-quality compost, how many days should you wait at least until you turn the material? *(open question; correct: 30; acceptable range 25 to 35)*

K7.) In order to produce good-quality compost, how many times should you turn the materials in the pit or heap until the composting is finished? *(open question; correct: 3)*

K8.) If you seed maize in lines, how wide should the distance between lines usually be?
(open question, assessed with measurement tape; correct: 75 to 80 cm; acceptable range: 65 to 90 cm)

K9.) If you seed faba beans in lines, how wide should the distance between lines usually be?
(open question, assessed with measurement tape; correct: 30 to 40 cm; acceptable range: 25 to 45 cm)

Principles Knowledge

K10.) For which purpose/benefit should you use improved seeds?
(open question; correct if mentions at least two correct points, i.e. one beyond "higher crop yield")

K11.) What are the major advantages of blended fertilizer (NPS+/NPK+) over DAP fertilizer? Which statements are correct?

K11_1.) Blended fertilizer contains a greater number of nutrients than DAP. *(correct)*
K11_2.) Nutrient supply is better balanced in blended fertilizer than in DAP. *(correct)*
K11_3.) Blended fertilizer directly improves soil structure.
K11_4.) Blended fertilizer is more suitable for your location's soil type than DAP. *(correct)*
K11_5.) Blended fertilizer controls weeds and pathogens.

K12.) Why is it important to use compost/organic fertilizer?
(open question; correct if mentions at least three correct points, i.e. two beyond "higher crop yield")

K13.) What are the major advantages of line seeding over broadcasting? Which statements are correct?

K13_1.) Line seeding reduces the crops' competition for space, nutrients and water. *(correct)*
K13_2.) Seeding a crop in lines is faster than broadcasting.
K13_3.) Line seeding reduces soil acidity.
K13_4.) With line seeding usually less seeds are needed. *(correct)*
K13_5.) With line seeding less fertilizer is needed because it can be targeted directly to the roots. *(correct)*
K13_6.) Line seeding makes weeding and harvesting easier. *(correct)*
K13_7.) Line seeding has no advantages.

K14.) Why is it important to use inorganic fertilizer and compost at the same time? Which statements are correct?

K14_1.) It is always better to apply inorganic fertilizer only.
K14_2.) Because the soil needs both organic and inorganic nutrient sources to be healthy and fertile. *(correct)*
K14_3.) Less seeds are needed when using inorganic and organic fertilizer at the same time.

K15.) What are the important characteristics of a fertile soil?

(open question; correct if mentions at least three correct points)

K16.) What are the benefits of applying inorganic fertilizer in lines or by band/microdosing? Which statements are correct?

K16_1.) It has no benefits.
K16_2.) It is faster than broadcasting.
K16_3.) It leads to less leaching of nutrients because they are directly targeted to the roots. *(correct)*

B2. Formula p-value correction

$$p_{adj} = 1 - (1 - p(k))^{g(k)}$$

Where $(k) = M^{(1-r(.k))}$, with

M as the number of tested outcomes in a family,

$r(.k)$ as mean correlation among all outcomes other than outcome k, and

$p(k)$ as the unadjusted p-value for the kth outcome.

Source: McKenzie (2012b), based on Sankoh et al. (1997) and used in Aker et al. (2011).

References

Aker, J. C., 2011. Dial “A” for agriculture: A review of information and communication technologies for agricultural extension in developing countries. Agricultural Economics 42: 631-647.

Aker, J. C., Boumnijel, R., McClelland, A. and N. Tierney, 2011. Zap it to me: The short-term impacts of a mobile cash-transfer program. Working Paper 268. Washington D.C.: Center for Global Development.

Anderson, J. R. and G. Feder, 2007. Agricultural Extension. In: Evenson, R. and P. Pingali (eds.): Handbook of Agricultural Economics 3: 2343-2378. Washington D.C.: The World Bank.

Angrist, J. D. and G. Imbens, 1994. Identification and estimation of Local Average Treatment Effects. Econometrica 62(2): 467-475.

ATA, 2014. National Strategy for Ethiopia’s Agricultural Extension System – Visions, Systemic Bottlenecks and Priority Interventions. Addis Abeba: Ethiopian Agricultural Transformation Agency (ATA).

Barrett, C. B. and L. E. M. Bevis, 2015. The self-reinforcing feedback between low soil fertility and chronic poverty. Perspectives. Nature Geoscience 8: 907-912.

Barrow, C. J., 1991. Land Degradation: Development and Breakdown of Terrestrial Environments. Cambridge University Press.

Belay, K., 2003. Agricultural extension in Ethiopia: The case of participatory demonstration and training extension system. Journal of Social Development in Africa 18 (1): 49-84.

BenYishay, A. and A. M. Mobarak, 2018. Social Learning and Incentives for Experimentation and Communication. Review of Economic Studies 0: 1-34.

Berg, G. and B. Zia, 2013. Harnessing Emotional Connections to Improve Financial Decisions – Evaluating the Impact of Financial Education in Mainstream Media. World Bank Policy Research Working Paper 6407. Washington D.C.: The World Bank.

Bernard, T., Makhija, S., Orkin, K., Taffesse, A. S. and D. J. Spielman, 2016. Video-Based Agricultural Extension – Analysis of a Pilot Project in Ethiopia. Project Note. Washington D.C.: IFPRI.

Bernard, T., Dercon, S., Orkin, K. and A. S. Taffesse, 2015. Will video kill the radio star? Assessing the potential of targeted exposure to role models through video. World Bank Economic Review 29 (1): 226-237.

Bernard, T., Dercon, S., Orkin, K., and A. S. Taffesse, 2014. The Future in Mind: Aspirations and Forward-looking Behaviour in Rural Ethiopia. Working Paper 2014-16. University of Oxford: Center for the Study of African Economies (CSAE).

Bernard, T., Dercon, S. and A. S. Taffesse, 2012. Beyond Fatalism – An Empirical Exploration of Self-efficacy Aspirations Failure in Ethiopia. Ethiopia Strategic Support Program Working Paper 46. Addis Abeba/Washington D.C.: Ethiopian Development Research Institute (EDRI)/IFPRI.

Cameron, A. C. and P. K. Trivedi, 2009. Microeconometrics using Stata. College Station/Texas: Stata Press.

Cameron, A. C. and P. K. Trivedi, 2005. Microeconometrics – Methods and Applications. Cambridge University Press.

Cappellari, L. and S. P. Jenkins, 2003. Multivariate probit regression using simulated maximum likelihood. The Stata Journal 3 (3): 278-294.

Chong, A., Duryea, S. and E. La Ferrara, 2012. Soap operas and fertility: Evidence from Brazil. American Economic Journal: Applied Economics 4 (4): 1-31.

Chong, A. and E. La Ferrara, 2009. Television and divorce: Evidence from Brazilian novelas. Journal of the European Economic Association 7 (2-3): 458-468.

CIA, 2018. The World Factbook: Ethiopia. Available: https://www.cia.gov/library/publications/the-world-factbook/geos/et.html, accessed: October 2018.

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

De Brauw, A., Eozenou, P., Gilligan, D. O., Hotz, C., Kumar, N. and J. V. Meenakshi, 2018. Biofortification, crop adoption and health information: Impact pathways in Mozambique and Uganda. American Journal of Agricultural Economics 100 (3): 906-930.

Duflo, E., Glennerster, R. and M. Kremer, 2007. Using Randomization in Development Economics Research: A Toolkit. Discussion Paper 6059. London: Centre for Economic Policy Research (CEPR).

FAO, 2016. FAOSTAT: Ethiopia. Available: http://www.fao.org/faostat/en/#country/238, accessed: October 2018.

Feder, G., Murgai, R. and J. B. Quizon, 2004. The acquisition and diffusion of knowledge: The case of pest management training in farmer field schools, Indonesia. Journal of Agricultural Economics 55 (2): 221-243.

Foster, A. D. and M. Rosenzweig, 1995. Learning by doing and learning from others: Human capital and technical change in agriculture. Journal of Political Economy 106 (3): 1176-1209.

Gandhi, R., Veeraraghavan, R., Toyama, K. and V. Ramprasad, 2009. Digital Green: Participatory video and mediated instruction for agricultural extension. Information Technologies & International Development 5 (1): 1-15.

Gautam, M., 2000. Agricultural extension: The Kenya experience – An impact evaluation. Washington D.C.: The World Bank.

Godtland, E. M., Sadoulet, E., De Janvry, A., Murgai, R. and O. Ortiz, 2004. The impact of farmer field schools on knowledge and productivity: A study of potato farmers in the Peruvian Andes. Economic Development and Cultural Change 53: 63-92.

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

Hicks, R. and D. Tingley, 2011. Causal mediation analysis. The Stata Journal 11(4): 605-619.

Imai, K., L. Keele, and D. Tingley, 2010. A general approach to causal mediation analysis. Psychological Methods 15: 309-334.

Imai, K., Keele, L., Tingley, D. and T. Yamamoto, 2011. Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. American Political Science Review 105(4): 765-789.

Jensen, R. and E. Oster, 2009. The power of TV: Cable television and women's status in India. The Quarterly Journal of Economics 124 (3): 1057-1094.

Knowler, D. and B. Bradshaw, 2007. Farmers' adoption of conservation agriculture: A review and synthesis of recent research. Food Policy 32: 25-48.

Kondylis, F., Mueller, V. and J. Zhu, 2017. Seeing is believing? Evidence from an extension network experiment. Journal of Development Economics 125: 1-20.

Kondylis, F., Mueller, V. and J. Zhu, 2015. Measuring agricultural knowledge and adoption. Agricultural Economics 46 (3): 449-462.

Laajaj, R. and K. Macours, 2017. Measuring skills in developing countries. Policy Research Working Paper 8000. Washington D.C.: The World Bank.

McKenzie, D., 2012a. Beyond baseline and follow-up: The case for more T in experiments. Journal of Development Economics 99: 210-221.

McKenzie, D., 2012b. Tools of the trade: A quick adjustment for multiple hypothesis testing. World Bank Development Impact Blog.
Available: http://blogs.worldbank.org/impactevaluations/tools-of-the-trade-a-quick-adjustment-for-multiple-hypothesis-testing, accessed: October 2018.

Niu, C. and C. Ragasa, 2018. Selective attention and information loss in the lab-to-farm knowledge chain: The case of Malawian agricultural extension programs. Agricultural Systems 165: 147-163.

Noltze, M., Schwarze, S. and Matin Qaim, 2012. Understanding the adoption of system technologies in smallholder agriculture: The system of rice intensification (SRI) in Timor Leste. Agricultural Systems 108: 64-73.

Ogutu, S., Fongar, A., Godecke, T., Jackering, L., Mwololo, H., Njuguna, M., Wollni, M. and M.Qaim, 2018. How to make farming and agricultural extension more nutrition-sensitive: evidence from a randomised controlled trail in Kenya. European Review of Agricultural Economics: 1-24, doi: 10.1093/erae/jby049.

Pamuk, H., Bulte, E. and A. A. Adekunle, 2014. Do decentralized innovation systems promote agricultural technology adoption? Experimental evidence from Africa. Food Policy 44: 227-236.

Place, F., Barrett, C. B., Freeman, H. A., Ramisch, J. J. and B. Vanlauwe, 2003. Prospects for integrated soil fertility management using organic and inorganic inputs: Evidence from smallholder African agricultural systems. Food Policy 28: 365-378.

Rogers, E. M, 1995. Diffusion of Innovations. 4th edition. New York: Free Press.

Rola, A. C., Jamias, S. B. and J. B. Quizon, 2002. Do farmer field school graduates retain and share what they learn? An investigation in Iloilo, Philippines. Journal of International Agricultural and Extension Education 9 (1): 65-76.

Sankoh, A. J., Huque, M. F. and S. D. Dubey, 1997. Some comments on frequently used multiple endpoint adjustment methods in clinical trials. Statistics in Medicine 16 (22): 2529-2542.

Schwartzstein, J., 2014. Selective attention and learning. Journal of the European Economic Association 12(6):1423-1452.

Sheahan, M. and C. B. Barrett, 2017. Ten striking facts about agricultural input use in Sub-Saharan Africa. Food Policy 67: 12-25.

Spielman, D. J., Byerlee, D., Alemu, D. and D. Kelemework, 2010. Policies to promote cereal intensification in Ethiopia: The search for appropriate public and private roles. Food Policy 35: 185-194.

Swanson, B. E., 2008. Global Review of Good Agricultural Extension and Advisory Service Practices. Rome: FAO.

Taffesse, A. S., Dorosh, P. and S. Asrat, 2011. Crop Production in Ethiopia: Regional Patterns and Trends. Washington D.C.: IFPRI.

Tripp, R., Wijeeratne, M. and V. H. Piyadasa, 2005. What should we expect from FFS? A Sri Lanka case study. World Development 33 (10): 1705-1720.

Van Campenhout, B., Vandevelde, S., Walukano, W. and P. van Asten, 2017. Agricultural extension messages using video on portable devices increased knowledge about seed selection, storage and handling among smallholder potato farmers in Southwestern Uganda. PLoS ONE 12(1): e0169557.

Van den Berg, H. and J. Jiggins, 2007. Investing in farmers – the impacts of farmer field schools in relation to integrated pest management. World Development 35 (4): 663-686.

Vanlauwe, B., Bationo, A., Chianu, J., Giller, K. E., Merckx, R., Mokwunye, U., Ohiokpehai, O., Pypers, P., Tabo, R., Shepherd, K. D., Smaling, E. M. A., Woomer, P. L. and N. Sanginga, 2010. Integrated soil fertility management: Operational definition and consequences for implementation and dissemination. Outlook on Agriculture 39 (1): 17-24.

Van Mele, P., 2006. Zooming-in Zooming-out: A novel method to scale up local innovations and sustainable technologies. International Journal of Agricultural Sustainability 4 (2): 131-142.

Vasilaky, K., Toyama, K., Baul, T. and D. Karlan, 2018. Learning Digitally: Evaluating the Impact of Farmer Training via Mediated Videos. Working Paper. Columbia University: Earth Institute and Department of Economics.

Ward, P. S., Bell, A., Droppelmann, K., and T. G. Benton, 2018. Early adoption of conservation agriculture practices: Understanding partial compliance in programs with multiple adoption decisions. Land Use Policy, 70: 27-37.

Zoussou, E., Van Mele, P., Vodouhe, S. D. and J. Wanvoeke, 2010. Women groups formed in response to public video screenings on rice processing in Benin. International Journal of Agricultural Sustainability 8 (4): 270-277.