The Effects of Participatory and Video-based Extension on the Adoption of Integrated Soil Fertility Practices – 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

Low agricultural productivity and severe levels of land degradation, coupled with high population growth rates, result in yield gaps and continue to threaten both food security and ecosystems in many African countries. The concept of *integrated soil fertility management* (ISFM) provides a promising approach to tackle soil degradation and increase productivity among smallholders in Africa. The core of ISFM is the integrated application of organic and inorganic soil amendments, accompanied by a general improvement of agronomic techniques. Yet, ISFM knowledge and its uptake among farmers are still low. At the same time, "bottom-up" extension models are gaining increased importance, as well as non-traditional forms of agricultural education. In this study we assess the effects of a decentralized, participatory extension program and an additional video intervention on the adoption of five ISFM practices – compost, blended fertilizer, improved seeds, line seeding and lime – among small-scale farmers in Ethiopia using a randomized control trial. Our sample consists of over 2,400 households spread over three Ethiopian highland regions. We find that the treatments induce ISFM adoption, in particular of compost, line seeding and lime, as well as gains in knowledge. We further find evidence for diffusion effects from farmers who actively participate in the extension activities to their non-active peers in treatment communities regarding increased adoption of ISFM practices at the household level. However, when it comes to integrated use of the practices at the plot level – which is particularly important given the synergistic effects of the technologies – the extension treatment does not affect non-active farmers in treated communities. Yet, our results imply that these "left out" farmers can be reached with the video intervention, successfully encouraging integrated adoption among them.

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

Despite considerable achievements in past decades, achieving stable food security still is one of the major challenges the global community has to face, which is underlined by recent upward trends in absolute numbers of undernourished people worldwide (FAO, 2018). This appears even more crucial in the light of on-going population growth, projected to be particularly substantial in Sub-Saharan Africa (UN, 2017). However, productivity levels in many African remain low and agricultural growth in the past was often attributed to an expansion in area rather than an intensification of production, resulting in severe threats for ecosystems and a depletion of the natural resource base. Climate change and increasing competition for land further exacerbate the pressure on the environment as well as food systems, and call for strategies that increase food production in a sustainable way on the same (or even lower) area of land (Godfray et al., 2010). One of the major bottlenecks to a sustainable intensification of agricultural productivity is the high level of 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 (FAO, 2013; Barrow, 1991). Soil degradation commonly goes along with a substantial loss of soil fertility which can result in severe yield deficits, particularly threatening for the food security of rural communities in developing countries (Barrow, 1991). Hence, there is an urgent need for solutions that restore soil fertility and increase yields without further compromising the natural resource base. The idea of *integrated soil fertility management* (ISFM) provides a promising concept to tackle soil degradation and increase productivity among smallholders in developing countries. The core of ISFM is the integrated application of inorganic and organic soil amendments (like compost, crop residues or manure), targeted to local conditions. Complementarily, ISFM aims on a general improvement of agronomic techniques, like line application of seeds and other inputs, and the introduction of practices such as crop rotation, cereal-legume intercropping, minimum tillage, mulching, residue retention, agroforestry or green manuring (Vanlauwe et al., 2010; Place et al., 2003). Yet, a central question is how farmers can be convinced to engage in ISFM activities, given that 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). This seems particularly problematic against the background of widespread poverty among small-scale farmers, which frequently lets them opt for immediate-return rather than long-term sustainable strategies (Bishaw, 2001). Furthermore, even if a

household adopts several practices or inputs simultaneously, this does often not occur at the plot level, leaving synergetic potential of their integrated use unexploited (Sheahan and Barrett, 2017).

Lack of information and knowledge is a frequently cited barrier to the adoption of new agricultural practices (Aker, 2011; Foster and Rosenzweig, 1995). Oftentimes, farmers are simply not aware of new techniques and their benefits, or do not possess the necessary skills, know-how or confidence in their abilities. Providing agricultural extension services to farmers can bridge this knowledge and capacity gap, as its overarching goal is knowledge transfer from researchers to farmers. Yet, 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 resource-poor and marginalized farmers (Aker, 2011; Anderson and Feder, 2007). Considering the fact that many developing countries struggle with constrained public budgets, the questions of effectiveness of extension – achieving maximum outreach under minimal costs – is particularly important. Especially in African countries, this gave rise to the introduction of decentralized approaches, where public extension agents only train selected farmers (often referred to as contact, lead or model farmers) to adopt new techniques on their farms. Exposure to on-farm demonstrations and well-trained contact farmers is supposed to spur broader adoption in a community through peer learning (Kondylis et al., 2017; Swanson, 2008; Gautam, 2000). Simultaneously, whereas extension in the past was rather perceived as a "top-down" process, it is now more frequently perceived as inclusive "bottom-up" strategy, focusing on participatory and experiential learning approaches. Oftentimes, these developments go along with a shift from a mere output growth perspective to a more holistic view on how to achieve productivity increases and sustainable use of natural resources at the same time, tailored to specific agro-ecological conditions (Swanson, 2008).

Yet, given the high policy relevance, there is comparatively little rigorous evidence on the impact of different extension programs on farmers' adoption behavior to date (Aker, 2011; Anderson and Feder, 2007). One stream of literature deals with the most common participatory extension approach Farmer Field Schools (FFS), providing mixed evidence. For instance, Feder et al. (2004a) find that knowledge gains among FFS graduates are modest and do not diffuse to neighboring farmers, while Feder et al. (2004b) do not find effects of FFS on yields or improved pest management. In contrast, Davis et al. (2012) report that FFS significantly increase incomes among participants in three East African countries, while Godtland et al. (2004) find FFS to be associated with (short-term) knowledge gains. Regarding diffusion effects, Rola et al. (2002) as well as Tripp et al. (2005) report that increased knowledge among

FFS farmers does hardly trickle down to their neighbors. In their review of several FFS studies related to integrated pest management, van den Berg and Jiggins (2007) conclude that achievements with regards to pesticide reduction among FFS graduates have been substantial in Asia. However, diffusion does not occur easily due to the complex nature of the information. Similarly, recent experimental research on decentralized extension concludes that the contact farmer approach does indeed encourage technology uptake among trained contact farmers, but this does little to spur adoption among other farmers in communities (Kondylis et al., 2017), or that diffusion highly depends on similarity between communicators and target farmers (BenYishay and Mobarak, 2018). Nevertheless, Krishnan and Patnam (2013) find that while extension agents matter in an early stage to create awareness for an innovation, learning from neighbors is more persistent over time. Though, insights from other studies on social learning underline that learning within social networks are complex and multifaceted processes, while probably depending considerably on the specific context as well as traits of a technology (e.g. Matuscke and Qaim, 2009; Bandiera and Rasul, 2006; Conley and Udry, 2001). Similarly, a randomized study on innovation platforms in eight African countries suggests that the success of participatory extension strongly depends on the technology studied and specific forms of social capital prevalent in communities (Pamuk et al., 2014). Yet, all in all, while there is a substantial body of experimental studies on agricultural technology adoption that concentrates on the effect of policy instruments and economic incentives (Carter et al., 2016; Carter et al, 2013; Duflo et al., 2011; Giné and Yang, 2009; Ashraf et al., 2009 – just to mention some), few focus on the effectiveness of extension, but find mixed evidence regarding both adoption and knowledge (De Brauw et al., 2018; Kondylis et al., 2017; BenYishay and Mobarak, 2018; Pamuk et al., 2014).

A further recent trend in the area of agricultural education is the increasing presence of non-traditional forms of information provision, for example via radio, mobile phones, TV or video (Aker, 2011). Previous research has shown that video as information delivery channel has the potential to increase motivation and induce behavioral changes in farming communities. In Ethiopia, the Ministry of Agriculture, in cooperation with the NGO Digital Green, has piloted an approach to use video as agricultural extension tool in selected communities with preliminary findings suggesting a strong potential to encourage technology adoption (Bernard et al., 2016). In line with that, Zoussou et al. (2010) and Van Mele (2006) find positive effects in Benin on both knowledge and adoption of improved rice parboiling technologies as response to village-based screenings of instructional videos. In a controlled study in 16 Indian villages, Gandhi et al. (2009) find that village-based screenings of short “how-to videos” on

new agricultural practices lead to an adoption rate seven times higher compared to classical extension only. However, these studies did not use randomized designs and therefore potentially suffer from selection bias, while dealing with relatively small sample sizes. Yet, in a large-scale experimental setting, Vasilaky et al. (2018) detect significant impacts of video training on adoption of SRI (System of Rice Intensification) technologies in India. Within a randomized experiment among potato farmers in Uganda, Van Campenhout et al. (2017) find that video-based information provision lead to gains in knowledge and, most interestingly, even seems to trigger latent knowledge in domains that were not explicitly mentioned in the videos. Moreover, psychological barriers have shown to negatively influence agricultural production decisions, in particular, a low internal *locus of control*, which describes the general feeling of being in control of one's lives and able to influence it, and closely related, low perceptions of *self-efficacy*, i.e. an individual's belief of personally possessing the ability to perform a certain behavior and achieve expected outcomes (Wuepper and Sauer, 2016; Bernard et al., 2012; Rogers, 1995; Bandura, 1977). Yet, exposure to "vicarious experiences" via role models has proven the potential to reduce these barriers. Studies from other fields also show that medial exposure to role models can be effective in inducing behavioral change, for example with regards to fertility (Jensen and Oster, 2009; Chong et al., 2012; Chong and La Ferrera, 2009) or financial decisions (Berg and Zia, 2013). An experimental study conducted in rural Ethiopia delivers interesting insights regarding the effect of a video intervention on aspirations and future-oriented behavior of poor people (Bernard et al., 2014). The authors find significant treatment effects on a series of psychological and behavioral indicators, driven by identification of the audience with the main characters featured in the videos. All in all, there is evidence that video can be an effective tool for inducing changes in perceptions, knowledge and behavior by providing relevant information and exposing the audience to potential role models to whom they can relate. However, previous studies from the agricultural area mainly revolved around instructional videos and were mostly conducted in non-experimental settings.

In this research, we aim to contribute to the current body of literature on agricultural learning by assessing the effectiveness of a decentralized, participatory extension approach as well as an additional video intervention on farmers' knowledge and adoption of ISFM practices using a randomized control trial design (RCT). The study is conducted within the Integrated Soil Fertility Management Project of the German Agency for International Cooperation in Ethiopia and focuses on a set of five practices: Compost, blended fertilizer, improved seeds, line seeding and lime. We find that our treatments induce ISFM adoption, in particular of compost, line seeding and lime, as well as gains in knowledge. We further find evidence

for diffusion effects from farmers who actively participate in the extension activities to their non-active peers in treatment communities regarding increased adoption of ISFM practices at the household level. However, when it comes to integrated use of the practices at the plot level, the extension treatment does not affect non-active farmers in treated communities. Yet, our results imply that these "left out" farmers can be reached with the video intervention, successfully encouraging integrated adoption among them.

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

2. Setting and Experimental Design

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 in order to involve them as active stakeholders in the technology development, prioritization and transfer process and strengthen 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 three Ethiopian highland regions: Amhara, Oromia and Tigray. The extension activities of the project have been introduced as RCT in order to assess the effectiveness of the chosen approach. The target population consists of small-scale farmers in the three regions; within each region, six Woredas (districts) are targeted. The ISFM+ project is a component of the broader Ethiopian Sustainable Land Management Programme (SLMP)[1] and only operates in Woredas where physical land rehabilitation measures (stabilization of hillsides, erosion control measures) have been successfully introduced by the SLMP. Building on the SLMP achievements, it is thought that smallholders now have the capacity to address issues of soil fertility and productivity increase more effectively. During the initial phase of the project from 2015 to 2018, the use of five of so-called "quickwin" technologies are promoted 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 is ought to 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. *Line seeding* is promoted to replace the common practice of broadcasting seeds. It should 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 soil pH value.

Overall, the ISFM+ consists of two branches. The first package of interventions focuses on capacity building among government agricultural advisory staff. This includes experts from provincial and Woreda Bureaus of Agriculture as well as development agents (DA), the extension officers at Kebele level, the lowest administrative unit in Ethiopia. The main pillar

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

of the ISFM+ is the extension intervention. 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 the following section.

2.2 Experimental Design

This study builds on an RCT design, using microwatersheds (mws) as units of randomization. The first intervention to be tested regarding its impact on smallholders' knowledge and adoption decision is the participatory extension treatment which forms the core of the ISFM+ project. In addition to the extension intervention, half of the treatment mws were assigned to a video intervention. This allows us to test the pure effect of the extension program (in the following referred to as T1, compared to the non-treated control group C), the combined effect of the extension-plus-video intervention (called T2, compared to C), as well as the additional effect of the video only (T2 compared to T1). The start of the extension intervention was accompanied by a baseline survey among 15 randomly selected households in all treatment and control mws.

Treatment Description and Implementation

The extension intervention introduced by the ISFM+ project consists of the following elements: 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. Each FREG is led by three model farmers (MFs) that were appointed during kick-off workshops in the initial phase of the project. The MFs are trained by public extension agents (DAs) on ISFM techniques. Some of the MFs or FREG members may be replaced from season to season, but this is not defined in a fixed way. The central activity of the FREGs and MFs is the establishment and maintenance of demonstration plots. These are on-farm trials on which the package of improved practices is applied, next to plots that are operated according to traditional farmer practices. Hence, differences in biomass production and yields resulting from common in contrast to improved 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, MFs and FREG members share and discuss their experience with the implementation of the new practices and the results regarding improved soil fertility and productivity. Field day activities are mainly targeted to FREG members, although

in some communities other farmers do also participate. The ISFM+ extension treatment mainly aims on creating awareness and know-how about ISFM through a knowledge sharing process from DAs to MFs and from MFs to FREG members. Through that entry point, information should diffuse to the broader population of farmers in the communities, as illustrated in figure 1. Hence, this model heavily relies on peer-to-peer learning. Since the launch of the ISFM+ project in 2015, activities are on-going, regionally aligned with the course of the main harvest cycle (meher season).

[figure 1 about here]

Through the extension intervention farmers are mainly supposed to gain awareness for and knowledge about *how* to implement ISFM practices. The video intervention should provide an additional stimulus to farmers by explaining them more about their background. More specifically, the film 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 order to provide a realistic overall view. These farmers 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, this section was different for the three regions. 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 (Bernard et al., 2016; 2015; 2014; BenYishay and Mobarak, 2018). Notwithstanding, all three versions strictly followed 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 sample Woredas 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 that take place in the soil – such as hydrological cycles, the "work" of roots, microorganisms and nutrients. Complex soil processes and the relationship between farming practices, soil fertility and improved yields are presented in a simplified way so that farmers gain a better understanding on *why* the techniques improve soil

fertility and yields, i.e. effectively enhance their "principles knowledge" (Rogers, 1995). The video screenings were conducted in T2 communities in early 2017 – around six weeks prior to the start of the main growing season – typically in farmers' training centers, health posts or schools, and followed by group discussions facilitated by DAs. Household heads were invited a few days prior to the screenings orally and with written invitation cards, in the case of double-headed households we invited both husband and wife. Only the 15 households from our baseline sample were invited and allowed to enter screening venues.

Sampling 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.[2] Microwatersheds are water catchment areas, i.e. natural hydrological entities defined by the topography of the land, which can consist of one or more communities that share a common rainwater outlet. 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/relatively little 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 (T2). In the second stage, in all treatment and control mws 15 households were randomly drawn from administrative lists to participate in the baseline survey. Figure 2 graphically depicts our full original sample, consisting of 2,416 households.

[figure 2 about here]

3. Data

In order to assess the interventions' impact, two rounds of survey data were collected among the 15 sampled farm households in each mws. A baseline survey took place in early 2016, shortly after the launch of the ISFM+ project in mid-2015. The timing of the survey may raise concerns that baseline data might be influenced by first project activities, which we cannot

[2] In particular, because the SLMP as umbrella project of the ISFM+ has equally been implemented at the mws level.

completely rule out. Yet, in the first months of operation, the project was still in the consolidation phase, i.e. conducting planning and kick-off workshops, while structured extension activities were implemented in very few communities yet. 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.

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

3.1 Survey Content

Both surveys covered modules on household sociodemographic characteristics, income and assets, food security level, social relationships, farming practices and agricultural production 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 questionnaire, 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 with questions on ISFM awareness, how-to and principles knowledge (see appendix B1).[3]

3.2 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

[3] Enumerators were intensively trained and supervised during a ten-day training period, both with respect to the content of the questions as well as the way of posing them, particularly relevant for the knowledge and psychological modules. Questionnaire contents were carefully translated into the three local languages Amharic, Afan Oromo and Tigrigna and pretested in several rounds.

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 main crops, and use improved seeds.[4] 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.

3.3 Key Outcome Variables

Our key outcome of concern is the adoption of ISFM quickwin practices, 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. In addition to the individual adoption decisions, we are interested in how many ISFM practices farmers adopted in the preceding harvest cycle, and therefore assess the *number of adopted quickwin technologies*, ranging from 0 to 5. We use this variable as our core measure 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. As further aggregated adoption measure we use a dummy variable taking 1 if the household *adopted the full quickwin package*, i.e. all four technologies, this time excluding lime (since adoption would otherwise always be zero in Tigray).[5] Moreover, it is important to perceive ISFM as an *integrated* approach, which means that practices are complementary to each other and need to be utilized together on the same plot in order to deploy their full potential with regards to increasing soil fertility and productivity. We assess this with a binary variable that measures whether a farmer has used all four practices in combination, i.e. *adopted the full quickwin package on the same plot* (again, without lime). Finally, in order to get a sense for adoption intensity, we evaluate the *share of land* on which farmers use the complete quickwin package, both *unconditional* and *conditional* on adoption.[6]

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

[5] 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, the conclusions we draw stay unaltered.

[6] Due to data availability, we assess this for the share of main crop land only – i.e. land planted with teff, wheat, barley, maize or sorghum.

Apart from adoption, we are also interested in the effect of our treatments on ISFM knowledge, as potential impact pathway to adoption. Following Rogers (1995), we distinguish between three types of knowledge: *awareness*, *how-to* and *principles* knowledge. For awareness knowledge, we first capture how many of the quickwin technologies farmers recall as ISFM technologies *by memory*, assessed with an open question. In a second step, we record which techniques respondents remember *by name*, by letting enumerators read through a list of practices.[7] We do this in order to distinguish between active and passive ISFM awareness, since an open-formatted recall question is assumed to require substantially more cognitive effort than the closed-formatted version. Subsequently, we administered a set of questions regarding the mode of implementation of the practices (how-to knowledge), as well as their purpose (principles knowledge). Farmers who were not aware of a practice in the first place, were immediately given a value of zero in the respective how-to and principles question. We varied the mode of asking between 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). For our knowledge variables, we construct scores for each domain as well as an overall knowledge indicator, which range from 0 to 1, with 1 standing for full awareness, respectively having correctly answered all questions in a domain. Although originally, few questions on other technologies were included in the exam, we only use quickwin relevant knowledge (excluding lime) in these scores.

Furthermore, we are interested in the impact the interventions have on two core psychological outcomes, which we believe might exert a positive influence on adoption. We rely on Bernard et al. (2012), who have developed and validated a simple, yet informative, measure for *locus of control* in the Ethiopian context, based on two binary indicators that appear to capture the degree to which people believe in destiny/luck vs. their own responsibility. Specifically, we asked:

For each of the following, please tell me which of the two propositions you most agree with:

1: "Each person is primarily responsible for his/her success or failure in life"
0: "One's success or failure in life is a matter of his/her destiny"

1: "To be successful, above all one needs to work very hard"
0: "To be successful, above all one needs to be lucky"

[7] 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.

We then combine the two factors in an additive way to a variable taking 0 for a low, 1 for medium, and 2 for high internal locus of control. To capture *self-efficacy*, we asked farmers which were the two major factors that determined how much income they generated from their farming activities in the past two years, following Wuepper and Sauer (2016). We then assigned scores for each factor, ranging from 1 (low self-efficacy) if the answer emphasized on external factors (such as weather, prices, slope of the land) to 3 (high self-efficacy) if the response was internally focused (e.g. "hard work", "perseverance", "I adopted", "I learned"), while ambiguous answers were coded with 2 (for example, "soil fertility", "production"). The two answers given were averaged and recoded to a combined indicator from 1 (low) to 5 (high).

4. Conceptual Framework and Identification Strategy

In what follows, we first outline the conceptual model guiding our analytical framework and then explain the estimation procedure used to assess the effects of our treatments.

4.1 Conceptual Model

Figure 3 illustrates a schematic representation of the conceptual model. As described earlier, in addition to the extension interventions at the mws level, another component of the ISFM+ project deals with strengthening supply chains for ISFM-relevant inputs like blended fertilizer, improved seeds and lime, as well as the training of staff in the agricultural advisory system. Yet, albeit these measures appear crucial in order to reduce market failures and certainly play a key role in the adoption of those ISFM practices that require additional inputs, our research focuses on the extension activities at the mws level, since they were randomly implemented. Further, given our sampling strategy, we do not expect these supply-side interventions to distort our impact assessment, since they were implemented at the Woreda level and can therefore be assumed to benefit treatment and control mws to a similar extent. However, as a consequence, we have reason to assume that adoption rates will also increase in control communities.

According to Rogers' (1995) theory of innovation diffusion, the first necessary stage to technology adoption is the accumulation of knowledge, which he divides into three different forms: awareness, knowing how ("how-to knowledge") and knowing why ("principles knowledge"). Firstly, we hypothesize that the extension treatment increases farmers' *awareness* for ISFM and its yield-enhancing potential in treatment communities due to demonstra-

tion plots, FREG formation and field days. We further suppose the extension activities to enhance farmers' knowledge on *why* the techniques are beneficial, although to a relatively low extent. Next, based on successful awareness rising, we expect farmers to seek information on *how* to apply ISFM techniques. We presume this process of *how-to knowledge* seeking can equally take place with or without technically knowing *why* ISFM works, as long as farmers are convinced *that* it works. Providing extension is found to rather aim at improving awareness and how-to knowledge, but oftentimes neglects its role in the provision of principles knowledge (Anderson and Feder, 2007; Rogers, 1995). In line with that, we expect *principles knowledge* information to dilute along the course of the underlying chain of knowledge transmission (cp. figure 1). Yet, it has been shown that individuals' "competence to decide whether or not to adopt" is facilitated by principles know-how due to enhanced capacity to appraising the effectiveness and consequences of adoption (Rogers, 1995: 166; Hanna et al., 2014). Among T2 farmers, in addition to increasing ISFM awareness, we hypothesize that exposure to the video intervention does lead to measurable increases in principles knowledge, since this was the in focus of the intervention. It is further assumed that the video intervention enhances farmers' feeling of *self-efficacy* and internal *locus of control* via identification with potential role models and thus, additionally encourages technology adoption. Consequently, we hypothesize that ISFM adoption rates at endline will be significantly higher in T1 compared to control communities, but largest among T2 households.

[figure 3 about here]

4.2 Identification Strategy

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

$$Y_{i1} = \alpha + \beta\, T_i1 + \delta\, T_i2 + \varepsilon_j + \epsilon_{ij} \qquad (1)$$

where Y_{i1} denotes the respective outcome variable for household or individual 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 has additionally been invited to a 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. We expect model farmers in our sample to artificially

bias results upwards, as they have been provided with direct training and inputs for the maintenance of demonstration plots and hence, exclude them from these analyses.

Although in theory, treatment indicators should be orthogonal to further explanatory variables, we will re-estimate all models including additional covariates in order to increase precision of our estimates – even more considering the initial imbalances in our sample (outlined in section 3.2):

$$Y_{i1} = \alpha + \beta\, T_i1 + \delta\, T_i2 + \gamma\, X_{i0} + \Omega\, \mathrm{W_j} + \nu\, Y_{i0} + \varepsilon_j + \epsilon_{ij} \quad (2)$$

In these models, X_{i0} represents a vector of control variables related to farmer and household characteristics captured at baseline, while with $\mathrm{W_j}$, 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, since we assume some degree of path dependency on previously gained experience with a technology, and thus reducing the overall variance.[8] This specification of treatment effect model has been shown to be more powerful than the difference-in-difference estimator in the presence of relatively low autocorrelation in the outcome measure (McKenzie, 2012a; De Brauw et al., 2018). Further, in each model we run a test of equality between β and δ, in order to evaluate the additional impact of the video intervention.

Our main parameters of interest β and δ give the intent-to-treat effect (ITT), which measures the average effect of being randomly assigned to T1 or T2, irrespective of actual participation. Since in reality compliance is never expected to be perfect, ITT estimates are of particular interest for policy makers. 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; Bloom, 2006). 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 (Duflo et al., 2007; Angrist and Imbens, 1994): First-

[8] 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 and psychological outcomes were not measured in the baseline survey.

ly, 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, adoption is affected via spillovers from actively participating individuals 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, active beneficiaries of the extension treatment that are involved in demonstration plot and field day activities. Consequently, since the majority of sample farmers in T1 and T2 are not expected to actively participate in the interventions, but rather profit indirectly via informational spillovers, we neither expect our treatment variables to be strongly correlated with active participation. 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 active FREG membership in the treatment groups:

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

where X_i denotes a vector of farmer and household covariates assumed to influence farmer $i's$ participation decision 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. 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. Nevertheless, are well aware that the chosen approach does not entirely eliminate self-selection bias, since matching only occurs on observable characteristics and hence, interpret the obtained results with caution and provide further robustness checks.

5. Results

In this section, we first present a descriptive overview on treatment participation as well as changes in our main outcome variables between baseline and endline for the different treatment groups. Subsequently, we present and discuss ITT results on the effects of our interventions.

5.1 Descriptives

Treatment participation

Among the two treatment groups, 84 of the sampled 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. Excluding model farmers, we further find that 120, or around 12% of T1 and T2 farmers, are active FREG members, meaning they formally belong to a group and have participated in field day activities along the course of the preceding season. In addition, 77 (8%) of 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 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.

Outcome measures

Table 1 provides a descriptive overview on our main outcome variables compared between treatment groups, measured at baseline (for those available) and endline. Regarding all aggregated measures (Panel A), it appears that at endline, adoption rates in both treatment groups are significantly higher than in the control group. Yet, despite the fact that on average, more T2 than T1 households adopt, this difference is not statistically significant. Besides, it becomes evident that adoption between baseline and endline increased considerably among all three groups, i.e. including control farmers. For example, whereas in baseline the number of

quickwin practices used was around or almost one in all three groups, we find that at endline, treatment farmers on average adopt almost three out of five practices and control farmers more than two. This confirms our assumption that the project's interventions at the Woreda level also affect farmers in control mws positively. If we look at adoption of individual ISFM practices (Panel B), we detect a similar pattern: among all groups, the use of the practices is more widespread in endline than baseline, but this increase is significantly larger among treated smallholders. Changes between baseline and endline seem particularly substantial for blended fertilizer (from an average adoption of 1.4% to 64%), and lime (from 1% to 13% on average), which have both been largely unavailable or unknown to farmers prior to the intervention, as well as improved seeds (from a mean adoption of 12% to 62%).[9] Further, descriptively we find indication for improvements in ISFM knowledge through our interventions (Panel C). Though we cannot provide comparisons with ex-ante knowledge, scores in all domain domains appear highest among T2 respondents, followed by T1 and C farmers, with all differences between being statistically significant, except for principles knowledge. Yet, descriptives provide rather mixed evidence regarding psychological outcome variables (Panel D): Although extension-plus-video farmers show higher values than extension-only farmers regarding locus of control, they do not so in comparison to the control group. Regarding self-efficacy, the opposite pattern is true, as T2 farmers score significantly higher than control farmers on this variable, but not compared to T1 farmers. Nevertheless, all in all, these descriptive results provide first evidence for treatment effects, which will be examined more thoroughly in the next section.

[table 1 about here]

5.2 Adoption Decision

Aggregated adoption measures

Table 2 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 four different regression specifications.[10] 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

[9] Though the latter was probably measured with noise during baseline and was likely underestimated.

[10] 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).

continuous or ordered process, we also estimate a linear as well as ordered probit and logit models in order to underline the robustness of our findings.

[table 2 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.[11] From the linear model we see that farmers in T1 adopt additional .593 practices, respectively .394 when including further covariates. Households in T2 appear to adopt .749 more practices than those in the control group (who on average adopt 2.217 practices), or .491 when controlling for baseline covariates (columns (1) and (2)). 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. Compared to control farmers, these estimates correspond with an increase in the number of adopted practices of 27%, respectively 17%, for farmers in T1, and 34%, respectively 22%, for farmers in T2. However, across all specifications, p-values of the tests for equality of T1 and T2 (.436, .259, .432, .287, columns (1) to (4)) indicate that the difference between the two treatment groups with regard to the number of adopted ISFM practices is not statistically significant.

Table 3 reports ITT results on two different measures for adoption of the full ISFM package as well as adoption intensity. For all specifications, we define adoption of the complete package as having adopted all four practices (compost, blended fertilizer, improved seeds, line seeding), excluding lime.[12] Columns (1) to (4) show main and AME of being assigned to T1 and T2 on the adoption of the complete quickwin package at the farm level using a probit regression. All results are statistically significant and indicate that farmers in T1 have a likelihood to adopt the full set of practices that is 11.5, respectively 8.8 (with additional controls), percentage points higher compared to control group farmers (who adopt in 19.1% of the cases), whereas smallholders in T2 are 16.2, respectively 12.2, percentage points more likely than control farmers (columns (3) and (4)). Yet, the difference in effect sizes between the two

[11] Since the coefficients of the ordered probit and logit 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 and logit models is violated which makes these results less reliable (Cameron and Trivedi, 2009).

[12] Yet, in table A3 (appendix A) we show that using the two alternative specifications of this measure does not alter the interpretation of our results.

treatment groups is not statistically significant (p-values of equality tests .407 and .272, columns (1) and (2)).

[table 3 about here]

As stated earlier, we also deem it necessary to consider whether households apply all four practices together on the same (main crop) plot.[13] Columns (5) to (8) present probit regression results for the adoption of the full quickwin package on at least one plot, which are positive and statistically significant. The AME show that households in T1 are 7.8, or 5.7 (with additional controls), percentage points more likely than control group households to adopt the full set on the same plot, while the likelihood for farmers in T2 is 10.8, respectively 8.3, percentage points higher than that of the control group, in which on average 15.3% of households adopt. However, again we do not detect a statistically significant difference between the effect sizes of T1 and T2 (p-values of equality tests .542 and .354, columns (5) and (6)).

As a measure of adoption intensity we consider the share of (main crop) land on which the full quickwin package was practiced, both conditional and unconditional on adoption, presented in columns (9) to (12). Again, we find positive and significant results of both interventions on smallholders' share of land under complete quickwin. Unconditionally, the estimates show that on average, households in the extension treatment dedicate around 3% more of their land to the full set of recommended practices than households in the control group, who use it on 3.9% of their land. For farmers in the extension-plus-video group this treatment effect is slightly higher, with 4.3%, respectively 3.4%, but this difference between T1 and T2 is again not statistically significant (p-values of equality tests .635 and .634, columns (9) and (10)). Columns (11) and (12) report treatment effects on the share of land under full ISFM conditional on adoption. Without controlling for further covariates, farmers in T1 appear to practice full ISFM on 6.5% more of their land than control farmers, while farmers in T2 seem to practice it on 5.4% more, with this differential treatment effect not being statistically significant (p-value of equality test .764, column (11)). Yet, these positive treatment effects do not appear robust to the inclusion of additional explanatory variables (column (12)).

Overall, it seems that both treatment arms, extension-only as well as extension-plus-video, exert a positive and significant influence on the decision to adopt ISFM, and these findings are robust to different specifications of the aggregated ISFM measure. Yet, we do not find evidence that the treatments have a robust positive effect on ISFM adoption intensity defined

[13] Here we only consider adoption on plots planted with main crops (teff, wheat, barley, maize or sorghum) due to data availability.

as share of land under full ISFM, once farmers have decided to adopt. Further, although effect sizes are consistently larger for T2 than for T1 (with the exception of the conditional share of land), these differences are not statistically significant, indicating that the offer to watch the movie did not provide an additional stimulus for adoption.

Individual adoption decisions

After having provided evidence of the positive and significant ITT effect of the interventions on the number of adopted ISFM technologies as well as the complete quickwin package, 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.[14, 15] Panel A in table 4 presents the effects 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, both in the specifications with and without further covariates. In contrast, we find no significant effect of the extension intervention on the adoption of blended fertilizer. For T2, we detect positive and significant treatment effects on all individual adoption decisions, albeit effects for blended fertilizer and improved seeds do not remain significant with the inclusion of additional explanatory variables.

[table 4 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 (Aker et al., 2011; 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

[14] 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.

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

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 (except in the specification for line seeding that includes controls, column (8)), 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 .298 and .198, columns (1) and (2); .526 and .980 columns (7) and (8); .294 and .106 columns (9) and (10)).

Panel B of table 4 depicts the average marginal effects of our estimates. For compost, the AME of T1 suggests that farmers in the extension treatment are 12.4 percentage points more likely to adopt than farmers in the control group, respectively 13.1 percentage points when adding further covariates. For T2, the AME indicate an increased likelihood of compost adoption of 19.3, or 17.1, percentage points, given that in the control group 40.4% of farmers adopt (columns (1) and (2)). Column (8) reports the AME for T1 and T2 regarding the adoption of line seeding (including controls), suggesting an increased likelihood to adopt of slightly more than 8 percentage points for both treatment arms. Columns (9) and (10) show the AME for lime adoption, indicating that farmers in Amhara and Oromia who are assigned to T1 are around 16 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 20 percentage points above the control group mean. These effects seem substantial, considering that in the control group on average 4% adopt.

Differential effects for FREG members

Until now, the ITT estimates represent the average effects of living in a community that has been assigned either to the extension-only (T1), or the extension-plus-video intervention (T2), irrespective of actual treatment status. Considering the video treatment, we assume 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 from model farmers – who are directly trained and provided with inputs – to FREG members during field day activities. Ideally, non-actively participating farmers in the communities will eventually be reached via informational spillovers. For our analyses, this means that a large proportion of farmers in in T1 and T2 have essentially not actively attended any extension activities. In fact, among the 987 farmers in T1 and T2, only 120 (12%) are active FREG members, i.e. the primary target group of the extension activities. Consequently, the remaining 847 farmers (88%) only benefit from the extension treatment through spillover effects; that is either via visiting a demonstration plot in their community independently, attending a field day or through information from their peers or development agent.[16] Hence, we are interested in whether the extension treatment has an effect on ISFM adoption beyond FREG membership, and whether the video treatment might influence FREG members and non-FREG members differently.

In order to get a first sense for the effects the interventions have for non-active beneficiaries, we re-estimate the ITT models excluding the 120 active FREG members from our sample (table A4, appendix A). Columns (1) to (6) indicate that both treatments exert a positive influence on the number of adopted ISFM practices as well as adopting the complete package at the farm level beyond active FREG membership. Yet, when considering the *integrated* use of the complete set of technologies on the same plot, T1 does not continue to exert a significant influence on adoption, signaling that the originally estimated effect regarding this outcome is concentrated among active FREG participants. In contrast, the coefficient of T2 stays positive and statistically significant in this subsample, indicating that T2 increases the likelihood to adopt the full package on the same plot by 9.3, or 7.8, percentage points compared to the control group, which is only slightly below the AME in the original model (table 3). In addition, the difference between the two point estimates is statistically significant in the specification with controls (p-value of equality test .059, column (8)). These results provide a first indication that the video intervention might have successfully conveyed the message that the *integrated* adoption of the practices is crucial to those farmers in treatment communities who do not actively benefit from the extension intervention and thus, spurs integrated adoption among this group of smallholders.

In a second step, we examine whether our treatments show differential effects for active FREG members by including two interaction terms between T1 and T2 with FREG member-

[16] In fact, recall that an additional 77 (9%) farmers in T1 and T2 group participated in a field day without formally belonging to a FREG, while 55 (6%) of the non-FREG treatment farmers state to have visited an ISFM demonstration plot independently, which may both be considered as (intended) spillovers. Moreover, 441 (51%) of these farmers state to have talked to the DA at least once in the past cropping season.

ship, respectively. Estimates in columns (1) and (6) of table A5 (appendix A) show that T1 seems to have a stronger effect for active FREG members regarding the number and adoption of the full set of ISFM practices, while the non-interacted treatment variable stays significant at the 10% level. Although we observe a similar pattern for T2, this differential effect appears to be weaker here.[17] Columns (7) and (8) of table A5 (appendix A) present the results of the interaction model regarding the joint adoption of all four practices on the same plot. Again, the video appears to offset some of the differential effect the extension intervention has for active participants when it comes to synergetic ISFM adoption, since T2 appears to work equally well for FREG as well as non-FREG farmers (insignificant interaction term, but significant main effect), whereas T1 does not seem to affect farmers beyond the subgroup of active beneficiaries (significant interaction term, but insignificant main effect), and (in the model with additional controls) the effects of T1 and T2 appear statistically different from each other (p-value of equality test .065, columns (8)).

However, when assessing the effect of FREG membership as core component of the extension treatment on technology adoption, the potential endogeneity problem becomes evident: The "better" farmers, i.e. the more progressive or skilled ones, might naturally adopt new technologies earlier or more intensively, while at the same time self-select into the FREGs. Hence, when simply using FREG membership as (interacted) explanatory variable in our regression framework or excluding FREG members from the analyses, results might be biased since they are potentially confounded with unobserved individual or household level characteristics. Yet, as previously discussed, we do not consider ToT estimates appropriate in our experimental design since we assume the extension treatment to have an effect beyond active participation.[18] In fact, with the two preceding estimation approaches we provide evidence for this effect, which eventually means the exclusion criterion of the instrumental variable method would not be fulfilled (Duflo et al., 2007; Angrist and Imbens, 1994). Therefore, in order to complement the previous findings, we use the PSM-inspired approach outlined in section 4.2.

[17] Yet, the Poisson model for the number of adopted practices with controls (column (4)) does not support the existence of differential effects for FREG members in T2.

[18] In fact, when using the instrumental variable approach to assess the ToT for active FREG members, we find these concerns confirmed, since estimates for the aggregated ISFM adoption measures appear implausible large. 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).

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.[19]

Table 5 reports that in both subsamples, treatment effects of the two interventions regarding the number of adopted ISFM practices remain positive and highly significant. 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, T1 seems to increase the number of applied technologies by .502, respectively .296, this coefficient rises to 1.178, or 1.071 with further covariates, in the FREG sample. Similarly, T2 is estimated to increase adoption by .680, or .437, practices in the non-FREG sample, but by 1.127, respectively 1.016, technologies in the FREG sample (OLS results in columns (1) to (4)).[20] A similar pattern becomes evident when considering the adoption of the complete package at the farm level, as columns (1) to (4) of table 6 show: Although the coefficients for T1 as well as T2 remain significant in the two subsamples, both treatment arms seem to spur adoption considerably more in the FREG than in the non-FREG sample. More precisely, regarding T1, the AME estimates indicate that within the non-FREG sample, treated, yet not actively participating, farmers are 8.6, respectively 5.7, percentage points more likely to adopt than their non-treated counterfactuals, while treatment farmers in the FREG sample appear to have an increased likelihood of 31.2, or 32.5, percentage points compared to hypothetical FREG farmers in the control group. For T2, the AME estimates in the non-FREG sample suggest an increased likelihood of adoption of 14.0, or 10.3, percentage points, whereas in the FREG sample these effects are estimated to over 29 percentage points. It becomes evident that the effect of T2 is almost twice as large for treated non-FREG farmers as that of T1, while the effects of the two treatments in the FREG sample are fairly similar. Finally, we examine the differential effects in the two subsamples regarding the integrated adoption of the technology package. Columns (5) to (6) 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 increases by around 22 percentage points in comparison to non-treated potential FREG farmers. Yet, the effect of T2 seems to behave in the opposite way. Somewhat sur-

[19] 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. Further, we find a high level of common support, since only two treated observations are off support. See appendices B3 and B4 for first-stage PSM regression results and a histogram of the estimated propensity score.

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

prisingly, this variable does not show a significant effect in the FREG sample, while in the non-FREG sample it appears to increase adoption by 10.2, or 7.9, percentage points.

[table 5 about here]

[table 6 about here]

We are confident to draw the following two conclusions: Firstly, the effect of the extension treatment is 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-active FREG members when it comes ISFM adoption at the household level. And secondly, our findings indicate that the video intervention compensates to some extent the effect of the extension treatment which is only absorbed via active FREG membership when it comes to *integrated* adoption, i.e. comprehending the importance of combining the practices on the same plot. This seems particularly important considering the general tendency among farmers in Africa, and in particular in Ethiopia, to simultaneously use organic and inorganic inputs at the household level, but fail to exploit synergy effects by combining them at the plot level (Sheahan and Barrett, 2017).

5.3 Impact Channels

The results of the previous section show that the experimental treatments influenced the adoption of ISFM quickwin technologies positively. As outlined earlier, the interventions were designed to foster adoption via the formation of different types of knowledge, and, somewhat subordinately, influencing farmers' perceptions of their own capacities and their general believes about being able to shape their livelihoods. We are therefore interested in examining whether increased ISFM adoption is indeed spurred via these hypothesized impact channels.

Knowledge

In table 7 we analyze the effect of the two interventions on the three different types of knowledge as well as a combined knowledge score.[21] All knowledge scores range from 0 to 1, with 1 meaning that a farmer is aware of all quickwin practices (for awareness knowledge), respectively answered all questions regarding their implementation (how-to knowledge) or purpose (principles knowledge) correctly. A score of 1 can therefore be interpreted as full, i.e.

[21] We cluster standard errors at the mws level, since this was the unit of randomization. Yet, as knowledge might also be correlated among spouses within one household due to intra-household knowledge sharing processes, we re-estimate all models clustering at the household level, which leads to similar results and does not alter implications (not reported).

100%, knowledge in the respective domain and coefficients as changes in percentage points. Panel A of table 7 reports the ITT estimates of T1 and T2 including both spouses' knowledge (in case of double-headed households), which we consider important since both had been invited to the video screenings. Panel B presents results for the sample of household heads only. Again, we apply the p-value correction for this family of knowledge outcomes in order to check robustness to multiple hypotheses testing.

[table 7 about here]

Columns (1) and (2) depict the results for "awareness by memory". From Panel A we learn that both T1 and T2 have a positive effect on recalling the quickwin practices by memory when the full sample is considered. While T1 farmers show a recall rate that is at least 3 percentage points above the control mean of 9.6%, T2 seems to increase this rate by at least 6 percentage points, with this effect sizes being statistically significantly different from each other (p-values of equality tests .021 and .023, columns (1) and (2)). If only the household heads are considered, T1 leads to an increase in awareness knowledge of over 5 percentage points, while that of T2 is around 8 (Panel B). Yet, in this sample, the differences between the coefficients of T1 and T2 are not significant (p-values of equality tests .113 and .124, columns (1) and (2)). When it comes to recognizing technologies by name among the sample including both spouses (Panel A), the coefficients of T1 with and without controls are .022, respectively .018, whereas that of T2 are .057, or .054, with the difference between them being significantly different at the 1% level (p-values of equality tests .003 and .001, columns (3) and (4)). Among household heads only (Panel B), awareness by name is estimated to increase by 2 percentage points through T1 in the model without additional covariates if we consider the naïve p-value, and does not show a significant effect if we add further explanatory variables to the model. In this specification, T2 appears to significantly increase the awareness-by-name score by over 0.04, compared to the control mean of .833. Again, the higher influence of T2 compared to T1 appears statistically different (p-values of equality tests .033 and .013, columns (3) and (4)). Yet, the coefficients of T1 turn insignificant in both specifications (with and without controls) within each sample if we apply the p-value adjustment, while that estimates for T2 remain highly statistically significant. Hence, the effect of the extension treatment on awareness knowledge is not robust to multiple testing, whereas the effect of the extension-plus-video treatment is.

Considering how-to knowledge, the two treatments carry positive and statistically significant coefficients in the models for both the full and the head-only samples. Up front, it ap-

pears that all results regarding this outcome stay significant at the 10% level after p-value correction. In the full sample (columns (5) and (6), Panel A), T1 seems to increase the share of correctly answered questions by at least 3 percentage points, whereas the effect of T2 is estimated to increase how-to knowledge by minimum 6 percentage points. Similarly, T1 is associated with a significant point estimate of .037, respectively .031, and T2 with a coefficient of .066, or .058, considering heads only (Panel B). In addition, the p-values of the tests for equality indicate that the larger effect of T1 over T2 is statistically significant in the full as well as in the head-only sample (p-values of equality tests .038 and .024 (Panel A), respectively .062 and .066 (Panel B), columns (5) and (6)).

Finally, columns (7) and (8) of table 7 report treatment effects on principles knowledge. In both samples, T1 is not estimated to exert a significant influence regarding this knowledge domain. In contrast, T2 is associated with somewhat modest, but statistically significant gains in principles knowledge in both subgroups in the specifications with controls, which are robust to p-value adjustment. Column (8) shows that T2 increases the share of correct responses by 4 (Panel A) and 3.1 (Panel B) percentage points, compared to 46.2% (both spouses) and 52.8% (heads only) correctly answered questions in the control group. Yet, while in the full sample T2 appears statistically different from T1 in the specification with controls, this effect is not present in the head-only sample (p-values of equality tests .042 (Panel A), respectively .190 (Panel B), columns (8)).

We can draw the following, preliminary conclusions from these results: When it comes to recalling ISFM technologies by memory, both treatments seem to exert a positive influence, while the effect size of T2 is significantly larger than that of T1 only in the sample with both spouses. This might be explained by the fact that female spouses in double-headed households are less frequently involved in agricultural trainings in general (Doss and Morris, 2001), as well as in the extension activities of our treatments. In fact, the large majority of field day participants in the treatment group (86%) states the head of their household attended solely, while in only 11% of double-headed households the female spouse attended. In contrast, out of all double-headed households in T2 that attended the screenings, in 80% of the cases the female spouse was present, which might contribute to the relatively stronger effect of the video intervention in the full compared to the household-head-only sample. Besides, we hypothesized that the video intervention significantly influences principles knowledge, since the intervention was particularly designed to explain *why* the technologies work, a knowledge domain we consider not being in the focus of the extension intervention. In fact, results show that extension-only does not increase farmers' principles knowledge, while farmers in the

extension-plus-video group show significantly higher background knowledge, both if we include spouses or not, although gains are somewhat lower than expected. Further, we obtain striking insights regarding the formation of how-to knowledge. Although both treatments lead to increases in this knowledge domain, it appears that the video intervention provides a significant effect in addition to the extension intervention. This is fairly surprising since the video did not explicitly convey information on *how* to apply the practices. Instead, it may be that it induced some information-seeking behavior on the mode of implementing ISFM.

Next, we are also interested in whether the treatments show differential effects on knowledge for active FREG members. Again, we include interaction terms between the two treatment variables and active FREG membership, as shown in table A6 (appendix A).[22, 23] We find that both treatments seem to increase awareness by memory relatively more for active FREG members, since the coefficients of both interaction terms are positive and statistically significant, while the significant main effects of both T1 and T2 suggest a treatment effect beyond FREG membership on memorizing ISFM technologies (columns (1) and (2)). Further, it appears that active FREG members in T2 benefit stronger than non-FREG members when it comes to awareness by name, although T2 still exerts a positive influence on this outcome beyond active FREG membership (columns (3) and (4)). In contrast, the extension-only treatment leads to increases in this knowledge outcome among active FREG members, while the main effect of T1 appears insignificant. Moreover, the p-values of the test of equality indicate that T2 increases awareness by name significantly more than T1 in these specifications (.052, .025, columns (3) and (4)). Columns (5) and (6) reveal that a similar pattern is true regarding how-to-knowledge: The effect of T1 seems to be concentrated among active beneficiaries and increases their how-to score by .152, or .109, but does not seem to influence non-FREG members significantly as the coefficient of T1 turns insignificant. In contrast, while T2 seems to enhance FREG members' how-to knowledge score by .120, or .082, it still carries a statistically significant coefficient for non-FREG members (of .051, respectively .048). In addition, the main effects of T1 and T2 appear significantly different from each other (p-values of equality tests .029 and .030, columns (5) and (6)). Lastly, it seems that T1 does not have an influence on principles knowledge, neither for FREG members nor other farmers, that

[22] In these models we only use the household heads' knowledge, since in all cases they are the household members that join the FREGs.

[23] In addition, we again apply the procedure outlined previously and re-estimate all knowledge regressions separately for the FREG- and non-FREG samples (not reported). Results of this approach lead to the same conclusions as the estimations with interaction terms.

is robust to additional control variables (columns (7) and (8)). For T2, the coefficients of the interaction variable with FREG participation are .126, or .075 if we add further covariates, and significant, signaling that T2 does have a positive effect for this subgroup. Yet, there does not seem to be a positive effect beyond active FREG membership that is robust to p-value adjustment.

Summing up, we have shown that both interventions have positive effects in different knowledge domains. However, gains in all three knowledge areas seem to be restricted to active FREG members in communities that have only been assigned to the extension treatment. Opposed to that, for T2 findings indicate robust treatment effects on all knowledge outcomes except principles knowledge also for non-FREG farmers, strengthening our previous argument that through the video intervention, farmers that are typically – or at least in our case – excluded from extension activities are reached and provided with ISFM-relevant knowledge. Nevertheless, extension activities alone seem to provide little knowledge spillovers from active beneficiaries to their peers within the same mws, although we have shown that adoption *is* positively influenced by the extension treatment beyond FREG membership. This is in line with previous literature showing that extension often does little in terms of (measurable) knowledge increases, but does still foster adoption (De Brauw et al., 2018; Kondylis et al., 2017; Feder et al., 2004a), which is shown to be oftentimes rather a result of imitation or herd effects (Krishnan and Patnam, 2013; Conley and Udry, 2001).

As a final step, we attempt to assess the relation between knowledge and adoption by including the knowledge scores as explanatory variables in the adoption regressions. Being aware of the potential endogeneity between knowledge and adoption, we interpret these results with caution, i.e. instead of establishing clear causality we are rather interested in the association between these variables. Table A7 (appendix A) depicts the effects of the how-to and principles knowledge scores on our main outcome variables.[24] It turns out that both knowledge scores positively influence ISFM adoption across all outcomes on a highly significant level. For example, possessing full knowledge in the how-to domain is associated with an increase in likelihood to adopt the complete ISFM package on the same plot by 43.3 percentage points, or 23.8 if we consider the specification with further controls, while full principles knowledge is related with an increase in this likelihood of 22.6, or 15.8, percentage points (columns (7)

[24] Since we perceive awareness of a practice as necessary precondition for adoption, we only include how-to and principles knowledge in these models. The correlation coefficient between these two knowledge variables is 0.3402, so multicollinearity is not expected to seriously bias these results.

and (8)). In table 8 we include both our treatment dummies as well as the two knowledge variables in the adoption models. Results continue to show a significant influence of knowledge on all adoption outcomes, while at the same time both T1 and T2 remain significant explanatory variables in this framework. Although their effect sizes drop compared to the models without knowledge covariates (cp. section 5.2, tables 2 and 3), the magnitude of these decreases cannot be considered substantial for any of the outcomes. We therefore conclude that large parts of the treatment effects cannot be attributed to gains in knowledge.

[table 8 about here]

Self-efficacy and Locus of control

Apart from providing smallholders with ISFM-relevant knowledge, the video intervention was designed to encourage adoption via socio-psychological pathways, in particular by enhancing farmers' feeling of self-efficacy and locus of control by exposing them to potential role models. First of all, we are therefore interested in the effects of the treatments on these outcomes, as illustrated in table 9.

[table 9 about here]

Since both outcomes are assessed as ordered variables, we estimate ordered probit in addition to linear models. It appears that T2 indeed shows a positive and significant influence on both locus of control as well as self-efficacy when the full sample is considered, albeit with relatively small effect sizes (Panel A). Yet, in the head-only sample, solely our measure of locus of control seems to be positively affected by T2, whereas it does not show a robust significant influence on self-efficacy (Panel B). In contrast, none of the estimated models reveals a significant effect of T1 on these variables. Moreover, regarding locus of control, coefficients for T2 are estimated to be larger than those of T1 on a statistically significant level (p-values of equality tests .032, .017, .036, .014 (Panel A) and .013, .007, .015, .006 (Panel B), columns (1) to (4)), leading to the conclusion that the video indeed affected this outcome. The somewhat remarkable finding that a movie of 20 minutes shown one year earlier induced measurable effect on this psychological outcome is well in line with the results of Bernard et al. (2014) who found similar impacts of a video intervention in rural Ethiopia on a variety of psychological and behavioral indicators. Yet, we cannot attribute a similar effect to the video intervention alone when it comes to self-efficacy (p-values of equality tests .636, .522, .635, .515 (Panel A), columns (5) to (8)), despite the significant coefficient of T2 identified in the

full sample, so we can only conclude that the combination of extension and video positively influences this measure.

Table A8 (appendix A) provides evidence on the relation between the household heads' locus of control as well as self-efficacy and ISFM adoption. We find no robust support for our hypothesis that increased locus of control might provide a stimulus for technology adoption. Regarding self-efficacy, evidence is rather week as well, since robustness of significant effects to the inclusion of further control variables is only given in the poisson specification for the number of adopted practices (column (4)), albeit with a fairly small AME of .056.

Ultimately, results shown in table 10 confirm these findings: when the variables locus of control and self-efficacy are added to the ITT models, effects of T1 and T2 stay very close to the originally estimated coefficients (cp. section 5.2, tables 2 and 3) while the self-efficacy variable continues to show only weak, and locus of control no significant influence at all. Hence, we can conclude that while the video treatment did in fact provide the intended effect on psychological measures at least to some extent, we fail to detect a clear-cut effect of these measures on ISFM adoption.

[table 10 about here]

6. Discussion and Conclusion

In this study we have assessed the impact of a decentralized, participatory extension program and an additional video intervention on the adoption of five ISFM practices – compost, blended fertilizer, improved seeds, line seeding and lime – among small-scale subsistence farmers in Ethiopia using an RCT design. We show that both treatment arms – extension-only as well as extension in combination with video – lead to increased adoption of ISFM technologies at the household level as well increased combined adoption on the same plot, which appears particularly important given the complementarities between the practices with regard to improvements in soil fertility and productivity. We fail to detect effects on adoption intensity measured as share of land under full ISFM, once farmers have decided to adopt. We find the main drivers of these treatment effects are enhanced adoption of compost, line seeding and lime, while impacts on adoption of blended fertilizer and improved seeds are not robust to adjusting p-values for multiple testing. This can be explained by the fact that farmers are probably less uncertain about the latter two 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. In addition, the use of improved seeds and mineral fertilizer is promoted by the overall advisory system and inputs are distributed via the Woreda Bureaus of Agriculture. Therefore, as expected, adoption in the control group also raised most significantly for these two practices. Though lime is also marketed by the Woreda Bureaus, this technology is largely new to farmers and received much more skeptically and therefore seems to require more informational inputs to create awareness for its benefits. Moreover, the use of blended fertilizer and improved seeds is to a large extent determined by financial constraints, since about 50% of respondents mention this as major obstacle to adoption for both practices. In contrast, the preparation of compost as well as seeding in lines present more labor intensive technologies – as over 60% respectively 45% of farmers mention labor or time constraints as primary barriers to adoption – while at the same time both practices are accompanied with substantial skepticism. For example, as learned during qualitative interviews, line seeding is often associated with lower yield expectations, since commonly less seeds and fertilizer are used when crops are planted in rows, while compost is sometimes seen to "suffocate" seeds. Hence, farmers need to be sufficiently convinced of their benefits so that the relative opportunity costs of time of alternative activities are reduced and they are willing to reallocate labor to these activities. Both the extension and the video intervention seem to have provided the necessary informational stimulus to alter farmers' yield expectations. However, prima facie, we do not detect any significant effect that can be purely attributed to the video regarding any aggregated or individual ISFM adoption measure, despite larger effect sizes of T2 over T1.

An important question of decentralized extension interventions revolves around the effectiveness of diffusion, i.e. whether there are beneficial effects for those who do not actively participate in activities. We find that active membership in a Farmer Research and Extension Group, the core component of the extension treatment, does indeed exert the strongest influence on ISFM adoption. Nevertheless, farmers residing in treatment communities still adopt significantly more quickwin technologies and more frequently adopt the full set of practices – measured at the farm level – even if they do not take part in FREG activities. 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-active treatment farmers, which is in line with results from previous studies reporting that exposure to adopting peers can hardly replace own experiential learning (Kondylis et al., 2017; Tripp et al., 2005; Feder et al., 2004a; Rola et al., 2002). However, the video intervention seems to offset this diffusion failure, since non-FREG members assigned to the additional video treatment do more frequently apply the package in an

integrated manner. This finding appears particularly important given that the behavior of farmers to spread different types of inputs across plots rather than exploiting their complementary potential is largely understudied (Sheahan and Barrett, 2017). In other words, the video appears to reach those farmers who are excluded from extension and bypassed by the information diffusion chain by successfully conveying the message of the synergistic effects of adoption.

We expected our treatments to spur adoption via two impact channels: first, increases in knowledge and second, two psychological pathways, namely locus of control and self-efficacy, shown to have influenced adoption in previous research. Regarding the latter two, we find evidence that the video indeed increased both psychological indicators at least partly, yet these increases contribute little to increase adoption. Regarding knowledge, our results suggest some impact of the extension treatment on ISFM awareness. Yet, we find a more clear-cut and significantly higher effect triggered through the video intervention. When it comes to knowing *why* ISFM is beneficial, we can neither report gains for FREG nor non-FREG farmers in the extension-only group. In contrast, for FREG members who were assigned to the additional video intervention, we find small but robust increases in principles knowledge, which points towards the complementary effects of extension and video regarding this knowledge domain. Interestingly, when it comes to knowledge about *how* to implement ISFM practices, the impact of T2 significantly exceeds that of T1, albeit no explicit how-to messages were conveyed in the video. Even more, extension-only does not show any influence beyond active FREG members regarding how-to knowledge. One possible explanation is that the video spurred how-to knowledge seeking processes, i.e. increased awareness for ISFM and its benefits encouraged farmers to gather information on the mode of implementation. 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 knowledge they already possess, even in areas not explicitly mentioned in the video. Besides, although ex-ante we hypothesized a stronger effect of the video on principles, but less on how-to knowledge, our finding make sense considering the lapse of time between video intervention and knowledge appraisal of one year. Whereas rather technical principles knowledge may have been formed as immediate response to the video intervention, it can be expected to fade over the course of time. In contrast, how-to knowledge needs to be readily available for the process of implementation, while at the same time gets constantly re-affirmed or updated through adoption itself (and hence, essentially making it an endogenous process). Further, we find that both principles and how-to knowledge are strongly associated with adoption, but

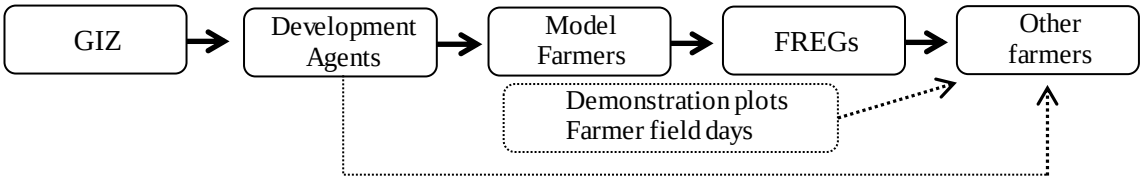
make rather small contributions in explaining the effects of our experimental treatments, which is in line with previous studies finding effects of extension on adoption that are only modestly explained by gains in knowledge (De Brauw et al., 2018; Kondylis et al., 2017). This might partly be attributed to potential measurement error knowledge assessments frequently suffer from, 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). Nevertheless, it is also plausible that knowledge constraints are simply not the most decisive barriers to adoption in our context. In contrast, our treatments might have played a more crucial role in influencing farmers' subjective yield expectations associated with ISFM, which has shown to be an important determinant to adoption in the literature (Knowler and Bradshaw, 2007). Besides, future research might focus on the relevance of practice- rather than package-specific knowledge for each individual technology, since information loss is known to occur due to selective attention to specific messages (Niu and Ragasa, 2018).

Another concern given our study setting is the extent to which control farmers were influenced by the experimental treatments, especially since we find adoption rates between baseline and endline have also risen in the comparison group. Firstly, we find that some, even if few, control farmers state to have participated in the extension interventions. Secondly, at least some of the studied technologies are currently getting incorporated in the Ethiopian extension system in general. And thirdly, we think higher level interventions of the ISFM+ project have positively affected control farmers as well by strengthening supply chains and expertise among agricultural advisory staff. In particular, we need to take into account that development agents in Ethiopia are assigned at the level of Kebeles, an administrative unit that can consist of one or several mws. Thus, training DAs on ISFM can also influence control mws that are located within the same Kebele as a treatment mws. Yet, when we split our sample into a potential spillover- and non-spillover sample – based on the criteria whether the mws is located in a Kebele in which a mws of opposite treatment status is present as well – and re-estimate our treatment models separately, we find no indication for differential effects. This lets us conclude that spillover effects resulting from direct communication between DAs and farmers are either not present, or not captured because they occur on an even higher than the Kebele level. Lastly, we can reasonably assume informational spillovers from treated to untreated farmers, potentially resulting in an underestimation of treatment effects. Assessing these hypothetical communication externalities in a more accurate way leaves us with potentially insightful opportunities for future investigation. Furthermore, more work needs to be done towards disentangling and quantifying the contributions the individual components of

the extension model – field day participation, community-based demonstration plots and FREG membership – as well as (perceived or actual) similarity between model and target farmers made to the success of the intervention, in order to a provide an even more solid basis for upscaling and transfer to other contexts. In addition, heterogeneous adoption patterns driven by regional and crop-specific factors should further be examined (Sheahan and Barrett, 2017). Although a replication of our interventions in other settings or with different populations may require adaptations, we feel confident with respect to the external validity of our findings given that the sample is spread over three regional states that in part differ quite substantially regarding agro-ecological, farming, climatic and cultural characteristics.

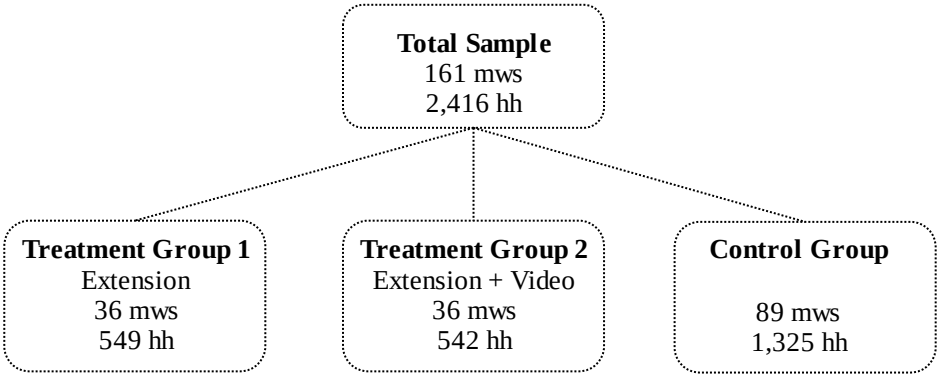
Tables and Figures

Figure 1. Intended information diffusion chain



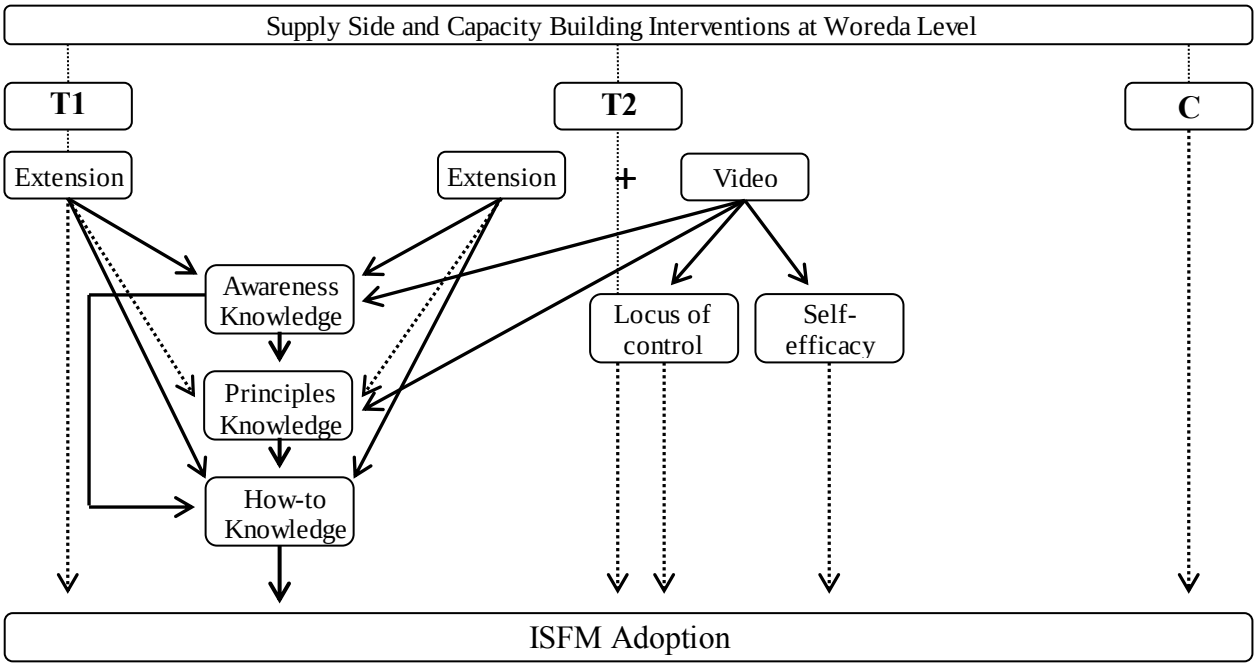
Source: Own illustration.

Figure 2. Diagrammatic illustration of the final sample



Source: Own illustration.

Figure 3. Conceptual model



Source: Own illustration.

Table 1. Descriptive overview on average baseline and endline outcomes by treatment group

	T1	T2	C	Overall	T1 vs. T2	T1 vs. C	T2 vs. C	N
Panel A: Aggregated Adoption Measures								
Number of quickwin technologies adopted (0 to 5)								
Baseline (proxy)	0.992	1.054	0.886	0.944	-0.062	0.106	0.168*	2,298
	(0.096)	(0.079)	(0.060)	(0.044)	(0.122)	(0.112)	(0.098)	
Endline	2.810	2.967	2.217	2.505	-0.157	0.593***	0.749***	2,298
	(0.144)	(0.142)	(0.111)	(0.081)	(0.200)	(0.180)	(0.179)	
Adoption of full quickwin (all 4, w/o lime)								
Baseline (proxy)	0.002	0.000	0.001	0.001	0.002	0.001	-0.001	2,298
	(0.002)	(0.000)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	
Endline	0.307	0.361	0.191	0.252	-0.054	0.116**	0.170***	2,298
	(0.044)	(0.049)	(0.027)	(0.022)	(0.066)	(0.052)	(0.056)	
Adoption of full quickwin on same plot (all 4, w/o lime)								
Endline	0.229	0.263	0.153	0.194	-0.034	0.077*	0.111**	2,076
	(0.038)	(0.043)	(0.024)	(0.018)	(0.057)	(0.045)	(0.049)	
Share of land under full quickwin (unconditional)								
Endline	0.074	0.082	0.039	0.056	-0.008	0.034**	0.042***	2,071
	(0.012)	(0.013)	(0.007)	(0.005)	(0.018)	(0.014)	(0.014)	
Share of land under full quickwin (conditional)								
Endline	0.323	0.312	0.258	0.291	0.011	0.065*	0.054**	401
	(0.031)	(0.019)	(0.017)	(0.013)	(0.036)	(0.035)	(0.026)	
Panel B: Individual Adoption Measures								
Adopted compost								
Baseline	0.325	0.357	0.365	0.355	-0.032	-0.041	-0.009	2,298
	(0.044)	(0.041)	(0.032)	(0.022)	(0.059)	(0.054)	(0.051)	
Endline	0.531	0.600	0.404	0.473	-0.069	0.126**	0.195***	2,298
	(0.048)	(0.047)	(0.035)	(0.025)	(0.066)	(0.059)	(0.058)	
Adopted blended fertilizer								
Baseline	0.010	0.021	0.014	0.014	-0.011	-0.004	0.007	2,298
	(0.005)	(0.009)	(0.005)	(0.004)	(0.010)	(0.007)	(0.010)	
Endline	0.691	0.703	0.595	0.639	-0.012	0.096	0.108**	2,298
	(0.051)	(0.042)	(0.035)	(0.025)	(0.066)	(0.061)	(0.055)	
Adopted improved seeds								
Baseline (proxy)	0.137	0.149	0.111	0.124	-0.013	0.026	0.039*	2,298
	(0.019)	(0.019)	(0.012)	(0.009)	(0.026)	(0.022)	(0.022)	
Endline	0.689	0.683	0.573	0.621	0.007	0.116*	0.110*	2,298
	(0.047)	(0.049)	(0.036)	(0.026)	(0.067)	(0.059)	(0.060)	
Adopted line seeding								
Baseline (proxy)	0.513	0.519	0.389	0.443	-0.006	0.124*	0.130*	2,298
	(0.057)	(0.062)	(0.036)	(0.028)	(0.084)	(0.067)	(0.072)	
Endline	0.772	0.809	0.622	0.695	-0.037	0.150***	0.187***	2,298
	(0.042)	(0.041)	(0.038)	(0.026)	(0.058)	(0.056)	(0.055)	
Adopted lime								
Baseline	0.012	0.009	0.012	0.011	0.003	0.000	-0.003	1,401
	(0.006)	(0.005)	(0.004)	(0.003)	(0.008)	(0.007)	(0.007)	
Endline	0.195	0.259	0.040	0.126	-0.065	0.155***	0.219***	1,401
	(0.039)	(0.050)	(0.011)	(0.018)	(0.063)	(0.040)	(0.051)	
Observations	505	482	1,311	2,298	987	1,816	1,793	

Panel C: Knowledge (endline)								
Overall	0.437	0.466	0.413	0.430	-0.029***	0.024**	0.052***	4,090
	(0.007)	(0.008)	(0.006)	(0.005)	(0.011)	(0.010)	(0.010)	
Awareness (memory)	0.126	0.158	0.096	0.116	-0.032**	0.030***	0.062***	4,090
	(0.009)	(0.010)	(0.005)	(0.005)	(0.014)	(0.011)	(0.011)	
Awareness (name)	0.826	0.861	0.804	0.821	-0.035***	0.022**	0.057***	4,090
	(0.008)	(0.008)	(0.007)	(0.005)	(0.011)	(0.011)	(0.011)	
How-to	0.359	0.389	0.323	0.345	-0.030**	0.036**	0.066***	4,090
	(0.010)	(0.010)	(0.009)	(0.006)	(0.015)	(0.014)	(0.014)	
Principles	0.471	0.492	0.462	0.470	-0.021	0.009	0.031*	4,090
	(0.012)	(0.013)	(0.009)	(0.006)	(0.018)	(0.015)	(0.016)	
Panel D: Psychological Factors (endline)								
Locus of control (0-2)	1.258	1.378	1.328	1.323	-0.120**	-0.070	0.051	4,090
	(0.040)	(0.039)	(0.025)	(0.019)	(0.056)	(0.047)	(0.047)	
Self efficacy (1-5)	1.906	1.930	1.840	1.874	-0.024	0.066	0.090**	3,966
	(0.038)	(0.034)	(0.021)	(0.016)	(0.051)	(0.043)	(0.040)	
Observations	905	864	2,324	4,093	1,769	3,229	3,188	

Note: Since line seeding and improved seeds were assessed slightly differently during baseline, baseline values present proxies of endline outcomes. For lime, we exclude households in Tigray since it is not recommended in this region and adoption is zero. Variables for adoption of the complete package and individual practices are dummies. Adoption at plot level is only measured for main crops (teff, wheat, barley, maize, sorghum). Missing values in the share of land variables stem from unidentified/unconvertible land measurement units; baseline values not available since line seeding was not assessed at plot level. All knowledge scores range from 0 (no knowledge) to 1 (full knowledge). Baseline values for knowledge and psychological variables not available, follow-up values include both spouses (in case of double-headed households). Model farmers excluded. 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 number of adopted quickwin technologies

	Number of quickwin technologies adopted							
	OLS		Poisson (AME)		Oprobit		Ologit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.593***	0.394***	0.593***	0.400***	0.467***	0.473***	0.741***	.833***
	(0.180)	(0.0728)	(0.173)	(0.0851)	(0.134)	(0.0773)	(0.236)	(0.139)
p-value	.001	.000	.000	.000	.001	.000	.002	.000
T2	0.749***	0.491***	0.729***	0.493***	0.607***	0.595***	0.977***	1.058***
	(0.179)	(0.0790)	(0.166)	(0.0841)	(0.139)	(0.0896)	(0.244)	(0.158)
p-value	.000	.000	.001	.000	.000	.000	.000	.000
Test T1=T2 (p-value)	.436	.259	.432	.287	.380	.215	.400	.180
Endline control mean	2.217							
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.056	.531	.012	.122	.020	.225	.017	.220
Observations	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298

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 DA, whether it has participated in agricultural training and grew main crops (teff, wheat, barley, maize, sorghum). Community level covariates are rainfall and temperature, whether Kebele 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 3. ITT effects on adoption of complete quickwin package at household and plot level, and adoption intensity

	Adoption of full quickwin				Adoption of full quickwin on same plot				Share of land under full quickwin			
	Main effects		AME		Main effects		AME		Unconditional		Conditional	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
T1	0.371**	0.402***	0.115**	0.088***	0.288*	0.288**	0.078*	0.057**	0.034**	0.028***	0.065*	0.014
	(0.160)	(0.123)	(0.0482)	(0.0257)	(0.158)	(0.116)	(0.0422)	(0.0228)	(0.0134)	(0.00861)	(0.0351)	(0.0290)
p-value	.020	.001	.017	.001	.069	.013	.064	.012	.012	.002	.067	.643
T2	0.520***	0.552***	0.162***	0.122***	0.398**	0.419***	0.108**	0.083***	0.043***	0.034***	0.054**	0.027
	(0.164)	(0.134)	(0.0486)	(0.0290)	(0.164)	(0.130)	(0.0435)	(0.0257)	(0.0144)	(0.0107)	(0.0257)	(0.0230)
p-value	.002	.000	.001	.000	.016	.001	.013	.001	.004	.002	.037	.236
Test T1=T2 (p-value)	.407	.272			.542	.354			.635	.634	.764	.624
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Endline control mean	.191				.153				.039		.258	
(Pseudo) R-squared	.025	.307			.015	.274			.017	.127	.019	.199
Observations	2,298	2,298	2,298	2,298	2,076	2,076	2,076	2,076	2,071	2,071	401	401

Note: Columns (1) to (8) show main and average marginal effects (AME) of probit models, columns (9) to (12) show OLS results (conditional and unconditional on adoption). All models show outcomes measured at endline, controlling for baseline level of the outcome (respectively a proxy) in one specification each (if available). Full quickwin package is a dummy variable defined as adopting all four practices, excluding lime. Reduced sample size for plot level indicator (columns (5) to (12)) because plot level data is only available for main crop plots. In the unconditional share indicator (columns (11) and (12)), plot sizes were given in non-convertible measurement units in five cases. Additional baseline control variables are the same as in table 2. 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 adoption of individual quickwin technologies

	Adopted compost		Adopted blended fertilizer		Adopted improved seeds		Adopted line seeding		Adopted lime	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Main effects										
T1	0.319**	0.524***	0.259	0.111	0.310*	0.248*	0.435**	0.401***	0.891***	1.119***
	(0.150)	(0.108)	(0.168)	(0.143)	(0.161)	(0.132)	(0.170)	(0.137)	(0.191)	(0.171)
Unadjusted p-values	.033	.000	.124	.435	.054	.060	.011	.003	.000	.000
Adjusted p-values	.103	.000	.357	.851	.187	.200	.040	.011	.000	.000
T2	0.495***	0.688***	0.294*	0.118	0.291*	0.180	0.563***	0.397***	1.107***	1.416***
	(0.150)	(0.106)	(0.151)	(0.138)	(0.164)	(0.149)	(0.179)	(0.144)	(0.202)	(0.204)
Unadjusted p-values	.001	.000	.052	.390	.075	.226	.002	.006	.000	.000
Adjusted p-values	.003	.000	.163	.808	.245	.604	.007	.022	.000	.000
Test T1=T2 (p-value)	.298	.198	.852	.966	.922	.678	.526	.980	.294	.106
Pseudo R-squared	.020	.360	.009	.226	.010	.250	.028	.398	.113	.323
Panel B: Average marginal effects										
T1	0.124**	0.131***	0.096	0.032	0.117*	0.070*	0.148***	0.084***	0.165***	0.157***
	(0.0572)	(0.0272)	(0.0618)	(0.0407)	(0.0599)	(0.0370)	(0.0573)	(0.0285)	(0.0364)	(0.0250)
T2	0.193***	0.171***	0.109*	0.034	0.110*	0.051	0.191***	0.083***	0.205***	0.199***
	(0.0559)	(0.0253)	(0.0558)	(0.0393)	(0.0610)	(0.0416)	(0.0598)	(0.0291)	(0.0395)	(0.0286)
Robust to adjustment?	Yes		No		No		Yes		Yes	
Endline control mean	.404		.595		.573		.622		.040	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298	1,401	1,401

Note: All models show main and 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 baseline control variables are the same as in table 2. 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 5. ITT effects on number of adopted quickwin technologies, FREG- and non-FREG samples separately

	Number of quickwin technologies adopted							
	OLS				Poisson (AME)			
	Non-FREG sample		FREG sample		Non-FREG sample		FREG sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.502***	0.296***	1.178***	1.071***	0.503***	0.302***	1.187***	1.079***
	(0.188)	(0.0745)	(0.217)	(0.161)	(0.181)	(0.0855)	(0.213)	(0.169)
p-value	.008	.000	.000	.000	.006	.000	.000	.000
T2	0.680***	0.437***	1.127***	1.016***	0.660***	0.440***	1.144***	1.111***
	(0.182)	(0.0771)	(0.242)	(0.223)	(0.168)	(0.0818)	(0.238)	(0.212)
p-value	.000	.000	.000	.000	.000	.000	.000	.000
Test T1=T2 (p-value)	.393	.102	.835	.797	.393	.115	.834	.820
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
Endline control mean	2.178		2.619		2.178		2.619	
(Pseudo) R-squared	.045	.533	.198	.559	.010	.126	.029	.861
Observations	2,062	2,062	236	236	2,062	2,062	236	236

Note: All models show outcomes measured at endline, controlling for baseline level of the outcome (respectively a proxy) in one specification each. Poisson models (columns (5) to (8)) show average marginal effects (AME). Number of quickwin technologies adopted ranges from 0 to 5. FREG- and non-FREG samples are constructed based on calculation of propensity scores predicting FREG membership. Additional baseline control variables are the same as in table 2. 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 adoption of complete quickwin package at household and plot level, FREG- and non-FREG samples separately

	Adoption of full quickwin				Adoption of full quickwin on same plot			
	Non-FREG sample		FREG sample		Non-FREG sample		FREG sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.086*	0.057**	0.312***	0.325***	0.055	0.029	0.220***	0.218***
	(0.0506)	(0.0257)	(0.0765)	(0.0652)	(0.0444)	(0.0215)	(0.0774)	(0.0634)
p-value	.089	.026	.000	.000	.213	.181	.004	.001
T2	0.140***	0.103***	0.296***	0.292***	0.102**	0.079***	0.123	0.092
	(0.0496)	(0.0282)	(0.0836)	(0.0658)	(0.0422)	(0.0249)	(0.0936)	(0.0725)
p-value	.005	.000	.000	.000	.016	.002	.188	.207
Test T1=T2 (p-value)	.361	.131	.880	.647	.355	.064	.321	.108
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
Endline control mean	.185		.246		.145		.218	
Pseudo R-squared	.019	.313	.081	.352	.013	.295	.035	.266
Observations	2,062	2,062	236	236	1,849	1,849	227	227

Note: All models show 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. Full quickwin package is a dummy variable defined as adopting all four practices, excluding lime. Reduced sample size for plot level indicator (columns (5) to (8)) because plot level data is only available for main crop plots. FREG- and non-FREG samples are constructed based on calculation of propensity scores predicting FREG membership. Additional baseline control variables are the same as in table 2. 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 effects on different knowledge outcomes

	ISFM Knowledge									
	Awareness				How-to		Principles		Overall	
	Known by memory		Known by name							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Both spouses										
T1	0.030***	0.031***	0.022**	0.018*	0.036**	0.030***	0.009	0.013	0.024**	0.022***
	(0.0107)	(0.0099)	(0.0111)	(0.0099)	(0.0139)	(0.0114)	(0.0150)	(0.0120)	(0.0096)	(0.0082)
Unadjusted p-values	.006	.002	.048	.079	.011	.010	.529	.278	.014	.007
Adjusted p-values	.012	.004	.105	.170	.025	.022	.826	.530		
T2	0.062***	0.060***	0.057***	0.054***	0.066***	0.060***	0.031*	0.040***	0.052***	0.053***
	(0.0112)	(0.0103)	(0.0108)	(0.0093)	(0.0137)	(0.0122)	(0.0157)	(0.0107)	(0.0098)	(0.0077)
Unadjusted p-values	.000	.000	.000	.000	.000	.000	.051	.000	.000	.000
Adjusted p-values	.000	.000	.000	.000	.000	.000	.114	.000		
Test T1=T2 (p-value)	.021	.023	.003	.001	.038	.024	.230	.042	.007	.001
Endline control mean	.096		.804		.323		.462		.413	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	.021	.084	.019	.111	.020	.153	.003	.233	.024	.277
Observations	4,090	4,089	4,090	4,089	4,090	4,089	4,090	4,089	4,090	4,089

Table 7 *(continued)*

Panel B: Household heads										
T1	0.054***	0.057***	0.020*	0.017	0.037***	0.031**	0.010	0.012	0.029***	0.027***
	(0.0135)	(0.0124)	(0.0113)	(0.0101)	(0.0141)	(0.0120)	(0.0159)	(0.0129)	(0.0096)	(0.0082)
Unadjusted p-values	.000	.000	.074	.103	.009	.010	.534	.365	.003	.001
Adjusted p-values	.000	.000	.161	.219	.020	.022	.834	.660		
T2	0.080***	0.079***	0.046***	0.043***	0.066***	0.058***	0.024	0.031**	0.052***	0.050***
	(0.0122)	(0.0108)	(0.0114)	(0.0097)	(0.0147)	(0.0128)	(0.0165)	(0.0119)	(0.0100)	(0.0074)
Unadjusted p-values	.000	.000	.000	.000	.000	.000	.146	.011	.000	.000
Adjusted p-values	.000	.000	.000	.000	.000	.000	.312	.026		
Test T1=T2 (p-value)	.113	.124	.033	.013	.062	.066	.455	.190	.031	.009
Endline control mean	.109		.833		.355		.528		.452	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	.033	.100	.014	.108	.020	.153	.002	.194	.025	.260
Observations	2,251	2,251	2,251	2,251	2,251	2,251	2,251	2,251	2,251	2,251

Note: All models show treatment effects on knowledge scores measured at endline, using OLS regressions. Knowledge scores range from 0 to 1 and were calculated based on the number of correct answers relative to the total number of questions in the respective domain. Knowledge exam was performed with both spouses in case of double-headed households. Additional baseline control variables are respondent's position in household (head or spouse), age, gender, education (in completed years), whether respondent had access to off-farm work or non-farm family business, whether household adopted the ISFM quickwin package at baseline and two regional dummies for Oromia and Amhara. 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 8. ITT and knowledge effects on main adoption outcomes

	Number of quickwin technologies adopted				Adoption of full quickwin		Adoption of full quickwin on same plot	
	OLS		Poisson (AME)		AME		AME	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.497***	0.362***	0.499***	0.367***	0.095**	0.081***	0.066*	0.047**
	(0.152)	(0.0704)	(0.147)	(0.0827)	(0.0431)	(0.0256)	(0.0392)	(0.0228)
p-value	.001	.000	.001	.000	.027	.002	.092	.039
T2	0.536***	0.419***	0.525***	0.418***	0.118***	0.104***	0.081**	0.069***
	(0.155)	(0.0756)	(0.144)	(0.0795)	(0.0446)	(0.0278)	(0.0409)	(0.0244)
p-value	.001	.000	.000	.000	.008	.000	.049	.005
How-to score	2.552***	0.937***	2.578***	1.087***	0.507***	0.225***	0.409***	0.219***
	(0.185)	(0.138)	(0.197)	(0.150)	(0.0539)	(0.0453)	(0.0523)	(0.0456)
p-value	.000	.000	.000	.000	.000	.000	.000	.000
Principles score	1.163***	0.779***	1.275***	0.907***	0.327***	0.182***	0.224***	0.148***
	(0.163)	(0.121)	(0.174)	(0.135)	(0.0529)	(0.0450)	(0.0499)	(0.0442)
p-value	.000	.000	.000	.000	.000	.000	.000	.000
Test T1=T2 (p-value)	.825	.491	.862	.543	.647	.416	.746	.414
Endline control mean	2.217				.191		.153	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.238	.557	.054	.129	.108	.326	.075	.293
Observations	2,251	2,251	2,251	2,251	2,251	2,251	2,033	2,033

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 (3) to (8)) show average marginal effects (AME). Number of quickwin technologies adopted ranges from 0 to 5, adoption of full quickwin package is a dummy variable defined as adopting all four practices, excluding lime. Reduced sample size for plot level indicator (columns (7) and (8)) because plot level data is only available for main crop plots. Knowledge scores range from 0 to 1 and were calculated based on the number of correct answers relative to the total number of questions in the respective domain; household heads' values used. Additional baseline control variables are the same as in table 2. 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 9. ITT effects on psychological outcomes

	Locus of Control				Self-efficacy			
	OLS		Oprobit		OLS		Oprobit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Both Spouses								
T1	-0.070	-0.048	-0.099	-0.074	0.132	0.088	0.138	0.096
	(0.0470)	(0.0434)	(0.0664)	(0.0639)	(0.0850)	(0.0635)	(0.0866)	(0.0676)
p-values	.140	.271	.136	.246	.122	.170	.110	.157
T2	0.051	0.069*	0.069	0.104*	0.181**	0.135**	0.187**	0.146**
	(0.0465)	(0.0358)	(0.0687)	(0.0552)	(0.0798)	(0.0568)	(0.0809)	(0.0604)
p-values	.276	.055	.312	.061	.025	.019	.021	.015
Test T1=T2 (p-value)	.032	.017	.036	.014	.636	.522	.635	.515
Endline control mean	1.328				2.680			
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.002	.063	.001	.033	.006	.082	.002	.030
Observations	4,090	4,089	4,090	4,089	3,966	3,965	3,966	3,965
Panel B: Head only								
T1	-0.060	-0.045	-0.087	-0.070	0.098	0.065	0.122	0.085
	(0.0516)	(0.0473)	(0.0744)	(0.0710)	(0.0696)	(0.0531)	(0.0854)	(0.0679)
p-values	.249	.346	.242	.324	.163	.220	.152	.208
T2	0.092*	0.106**	0.134*	0.166**	0.108	0.072	0.136*	0.096
	(0.0477)	(0.0409)	(0.0752)	(0.0680)	(0.0658)	(0.0494)	(0.0813)	(0.0637)
p-values	.055	.011	.074	.014	.104	.148	.095	.134
Test T1=T2 (p-value)	.013	.007	.015	.006	.904	.917	.893	.895
Endline control mean	1.376				2.702			
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.004	.070	.002	.038	.003	.072	.001	.032
Observations	2,251	2,251	2,251	2,251	2,209	2,209	2,209	2,209

Note: All models show treatment effects measured at endline, using OLS and ordered probit regressions. Indicator for locus of control ranges from 0 to 2, for self-efficacy from 1 to 5; both variables were assessed among both spouses in case of double-headed households. Additional baseline control variables are respondent's position in household (head or spouse), age, gender, education (in completed years), whether respondent had access to off-farm work or non-farm family business 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 10. ITT and psychological effects on main adoption outcomes

	Number of quickwin technologies adopted				Adoption of full quickwin		Adoption of full quickwin on same plot	
	OLS		Poisson (AME)		AME		AME	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.583***	0.390***	0.585***	0.400***	0.116**	0.091***	0.078*	0.057**
	(0.173)	(0.0764)	(0.167)	(0.0888)	(0.0479)	(0.0269)	(0.0420)	(0.0238)
p-value	.001	.000	.000	.000	.015	.001	.064	.017
T2	0.745***	0.503***	0.724***	0.510***	0.159***	0.125***	0.102**	0.081***
	(0.173)	(0.0814)	(0.160)	(0.0857)	(0.0483)	(0.0293)	(0.0431)	(0.0260)
p-value	.000	.000	.000	.000	.001	.000	.018	.002
Locus of control	0.053	-0.005	0.054	-0.032	0.030***	0.0002	0.017	0.005
	(0.0385)	(0.0298)	(0.0392)	(0.0331)	(0.0117)	(0.00982)	(0.0119)	(0.0107)
p-value	.169	.870	.166	.327	.010	.984	.143	.625
Self efficacy	0.220***	0.035	0.220***	0.048*	0.040***	0.004	0.037***	0.010
	(0.0398)	(0.0260)	(0.0392)	(0.0282)	(0.0128)	(0.00920)	(0.0119)	(0.00922)
p-value	.000	.179	.000	.089	.002	.626	.002	.298
Test T1=T2 (p-value)	.393	0.208	.394	.222	.441	.276	.610	.402
Endline control mean	2.217				.191		.153	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.079	.529	.017	.119	.034	.306	.023	.274
Observations	2,209	2,209	2,209	2,209	2,209	2,209	2,012	2,012

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 (3) to (8)) show average marginal effects (AME). Number of quickwin technologies adopted ranges from 0 to 5, adoption of full quickwin package is a dummy variable defined as adopting all four practices, excluding lime. Reduced sample size for plot level indicator (columns (7) and (8)) because plot level data is only available for main crop plots. Locus of control variable ranges from 0 to 2, self-efficacy from 1 to 5; household heads' values used. Additional baseline control variables are the same as in table 2. 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.

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 *(dummy)*	0.187	0.174	0.159	**0.169**	0.013	0.028	0.015	2,382

	(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 quickwin technologies adopted, using alternative 0-4 measure

	Number of quickwin technologies adopted							
	OLS		Poisson (AME)		Oprobit		Ologit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.489***	0.300***	0.488***	0.305***	0.396***	0.367***	0.651***	0.674***
	(0.170)	(0.0763)	(0.164)	(0.0877)	(0.138)	(0.0895)	(0.235)	(0.159)
p-value	.004	.000	.003	.001	.004	.000	.006	.000
T2	0.600***	0.359***	0.587***	0.366***	0.503***	0.445***	0.826***	0.814***
	(0.163)	(0.0777)	(0.154)	(0.0856)	(0.139)	(0.0946)	(0.237)	(0.169)
p-value	.000	.000	.000	.000	.000	.000	.000	.000
Test T1=T2 (p-value)	.532	.494	.531	.497	.491	.457	.508	.445
Endline control mean				2.195				
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.041	.509	.009	.110	.014	.228	.014	.221
Observations	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298

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 baseline control variables are the same as in table 2. 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 adoption of complete quickwin package, using alternative measures

	Adoption of full quickwin (at least 4 out of 5)				Adoption of full quickwin (region-specific)			
	Main effects		AME		Main effects		AME	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.461***	0.519***	0.149***	0.118***	0.492***	0.631***	0.078***	0.077***
	(0.152)	(0.106)	(0.0473)	(0.0231)	(0.155)	(0.135)	(0.0232)	(0.0156)
p-value	.002	.000	.002	.000	.002	.000	.001	.000
T2	0.624***	0.686***	0.203***	0.156***	0.681***	0.766***	0.108***	0.095***
	(0.156)	(0.124)	(0.0473)	(0.0271)	(0.170)	(0.151)	(0.0266)	(0.0182)
p-value	.000	.000	.000	.000	.000	.000	.000	.000
Test T1=T2 (p-value)	.336	.167			.243	.356		
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
Endline control mean			.202				.051	
(Pseudo) R-squared	.036	.326			.047	.274		
Observations	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298

Note: All models show main and 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 quickwin 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 baseline control variables are the same as in table 2. 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 main adoption outcomes, excluding active FREG members

	Number of quickwin technologies adopted				Adoption of full quickwin		Adoption of full quickwin on same plot	
	OLS		Poisson (AME)		AME		AME	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.459**	0.303***	0.458**	0.307***	0.0789	0.0591**	0.046	0.025
	(0.187)	(0.0743)	(0.180)	(0.0849)	(0.0505)	(0.0261)	(0.0443)	(0.0222)
p-value	.015	.000	.011	.000	.119	.024	.299	.256
T2	0.638***	0.446***	0.616***	0.448***	0.133***	0.107***	0.093**	0.078***
	(0.181)	(0.0766)	(0.167)	(0.0805)	(0.0495)	(0.0283)	(0.0422)	(0.0249)
p-value	.001	.000	.000	.000	.007	.000	.028	.002
Test T1=T2 (p-value)	.392	.095	.392	.106	.358	.117	.350	.059
Endline control mean	2.217				.191		.153	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.038	.530	.008	.124	.016	.313	.010	.298
Observations	2,178	2,178	2,178	2,178	2,178	2,178	1,957	1,957

Note: All models show outcomes measured at endline, controlling for baseline level of the outcome (respectively a proxy) in one specification each. Active FREG members excluded in these estimations. Poisson and probit models (columns (3) to (8)) show average marginal effects (AME). Number of quickwin technologies adopted ranges from 0 to 5, adoption of full quickwin package is a dummy variable defined as adopting all four practices, excluding lime. Reduced sample size for plot level indicator (columns (7) and (8)) because plot level data is only available for main crop plots. Additional baseline control variables are the same as in table 2. 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 main adoption outcomes, including heterogeneous effects for active FREG members

	Number of quickwin technologies adopted				Adoption of full quickwin		Adoption of full quickwin on same plot	
	OLS		Poisson (AME)		AME		AME	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.459**	0.307***	0.471**	0.328***	0.080	0.063**	0.047	0.025
	(0.187)	(0.0751)	(0.185)	(0.0878)	(0.0513)	(0.0266)	(0.0456)	(0.0230)
p-value	.015	.000	.011	.000	.119	.019	.299	.273
T1 x Active FREG	1.124***	0.747***	0.878***	0.491***	0.252***	0.169***	0.199***	0.176***
	(0.163)	(0.170)	(0.130)	(0.154)	(0.0597)	(0.0464)	(0.0512)	(0.0449)
p-value	.000	.000	.000	.001	.000	.000	.000	.000
T2	0.638***	0.447***	0.633***	0.471***	0.135***	0.108***	0.096**	0.077***
	(0.181)	(0.0771)	(0.172)	(0.0818)	(0.0503)	(0.0290)	(0.0436)	(0.0257)
p-value	.001	.000	.000	.000	.007	.000	.028	.003
T2 x Active FREG	0.895***	0.373*	0.683***	0.147	0.185***	0.085*	0.076	0.026
	(0.185)	(0.201)	(0.137)	(0.192)	(0.0659)	(0.0507)	(0.0547)	(0.0418)
p-value	.000	.066	.000	.444	.005	.096	.162	.530
Test T1=T2 (p-value)	.392	.103	.392	.097	.358	.146	.350	.065
Endline control mean	2.217				.191		.153	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.080	.539	.017	.123	.038	.314	.025	.285
Observations	2,298	2,298	2,298	2,298	2,298	2,298	2,076	2,076

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 (3) to (8)) show average marginal effects (AME). Number of quickwin technologies adopted ranges from 0 to 5, adoption of full quickwin package is a dummy variable defined as adopting all four practices, excluding lime. Reduced sample size for plot level indicator (columns (7) and (8)), because plot level data is only available for main crop plots. Additional baseline control variables are the same as in table 2. 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. ITT effects on different knowledge outcomes, including heterogeneous effects for active FREG members

	ISFM Knowledge									
	Awareness				How-to		Principles		Overall	
	Known by memory		Known by name							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
T1	0.034**	0.041***	0.013	0.012	0.019	0.018	-0.001	0.007	0.014	0.018**
	(0.0136)	(0.0125)	(0.0114)	(0.0102)	(0.0139)	(0.0119)	(0.0156)	(0.0127)	(0.0095)	(0.0082)
Unadjusted p-values	.013	.001	.261	.233	.173	.124	.934	.571	.137	.032
Adjusted p-values	.028	.002	.498	.453	.349	.259	.998	.866		
T1 x Active FREG	0.166***	0.138***	0.063***	0.037**	0.152***	0.109***	0.093***	0.039	0.119***	0.079***
	(0.0404)	(0.0397)	(0.0161)	(0.0142)	(0.0238)	(0.0273)	(0.0325)	(0.0274)	(0.0161)	(0.0152)
Unadjusted p-values	.000	.001	.000	.010	.000	.000	.005	.153	.000	.000
Adjusted p-values	.000	.002	.000	.023	.000	.000	.012	.326		
T2	0.061***	0.064***	0.037***	0.037***	0.051***	0.048***	0.0081	0.022*	0.037***	0.041***
	(0.0126)	(0.0114)	(0.0115)	(0.0100)	(0.0141)	(0.0124)	(0.0155)	(0.0118)	(0.0093)	(0.0072)
Unadjusted p-values	.000	.000	.002	.000	.000	.000	.604	.070	.000	.000
Adjusted p-values	.000	.000	.005	.000	.000	.000	.889	.158		
T2 x Active FREG	0.147***	0.127***	0.076***	0.051***	0.120***	0.082***	0.126***	0.075**	0.119***	0.082***
	(0.0329)	(0.0342)	(0.0168)	(0.0161)	(0.0196)	(0.0206)	(0.0286)	(0.0304)	(0.0153)	(0.0158)
Unadjusted p-values	.000	.000	.000	.002	.000	.000	.000	.014	.000	.000
Adjusted p-values	.000	.000	.000	.005	.000	.000	.000	.033		
Test T1=T2 (p-value)	.102	.107	.052	.025	.029	.030	.600	.307	.025	.008
Endline control mean	.109		.833		.355		.528		.452	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	.065	.122	.024	.112	.044	.164	.013	.198	.061	.276
Observations	2,251	2,251	2,251	2,251	2,251	2,251	2,251	2,251	2,251	2,251

Note: All models show treatment effects on household heads' knowledge scores measured at endline, using OLS regressions. Knowledge scores range from 0 to 1 and were calculated based on the number of correct answers relative to the total number of questions in the respective domain. Additional baseline control variables are respondent's position in household (head or spouse), age, gender, education (in completed years), whether respondent had access to off-farm work or non-farm family business, whether household adopted the ISFM quickwin package at baseline and two regional dummies for Oromia and Amhara. 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 A7. Knowledge effects on main adoption outcomes

	Number of quickwin technologies adopted				Adoption of full quickwin		Adoption of full quickwin on same plot	
	OLS		Poisson (AME)		AME		AME	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
How-to score	2.732***	1.053***	2.742***	1.196***	0.544***	0.254***	0.433***	0.238***
	(0.203)	(0.142)	(0.212)	(0.156)	(0.0550)	(0.0467)	(0.0535)	(0.0468)
p-value	.000	.000	.000	.000	.000	.000	.000	.000
Principles score	1.183***	0.818***	1.300***	0.952***	0.332***	0.198***	0.226***	0.158***
	(0.169)	(0.126)	(0.179)	(0.140)	(0.0538)	(0.0465)	(0.0509)	(0.0449)
p-value	.000	.000	.000	.000	.000	.000	.000	.000
Endline control mean	2.217				.191		.153	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.206	.541	.047	.126	.093	.313	.066	.286
Observations	2,251	2,251	2,251	2,251	2,251	2,251	2,033	2,033

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 (3) to (8)) show average marginal effects (AME). Number of quickwin technologies adopted ranges from 0 to 5, adoption of full quickwin package is a dummy variable defined as adopting all four practices, excluding lime. Reduced sample size for plot level indicator (columns (7) and (8)), because plot level data is only available for main crop plots. Knowledge scores range from 0 to 1 and were calculated based on the number of correct answers relative to the total number of questions in the respective domain; household heads' values used. Additional baseline control variables are the same as in table 2. 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 A8. Psychological effects on main adoption outcomes

	Number of quickwin technologies adopted				Adoption of full quickwin		Adoption of full quickwin on same plot	
	OLS		Poisson (AME)		AME		AME	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Locus of control	0.060	0.002	0.061	-0.022	0.032***	0.004	0.018	0.008
	(0.0387)	(0.0302)	(0.0395)	(0.0333)	(0.0116)	(0.00989)	(0.0123)	(0.0110)
p-value	.124	.950	.123	.514	.005	.682	.138	.479
Self efficacy	0.242***	0.043	0.240***	0.056*	0.045***	0.007	0.041***	0.012
	(0.0427)	(0.0266)	(0.0419)	(0.0290)	(0.0134)	(0.00930)	(0.0120)	(0.00912)
p-value	.000	.108	.000	.053	.001	.473	.001	.181
Endline control mean	2.217				.191		.153	
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
(Pseudo) R-squared	.024	.507	.005	.114	.010	.287	.009	.2646
Observations	2,209	2,209	2,209	2,209	2,209	2,209	2,012	2,012

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 (3) to (8)) show average marginal effects (AME). Number of quickwin technologies adopted ranges from 0 to 5, adoption of full quickwin package is a dummy variable defined as adopting all four practices, excluding lime. Reduced sample size for plot level indicator (columns (7) and (8)) because plot level data is only available for main crop plots. Locus of control variable ranges from 0 to 2, self-efficacy from 1 to 5; household heads' values used. Additional baseline control variables are the same as in table 2. 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.

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.

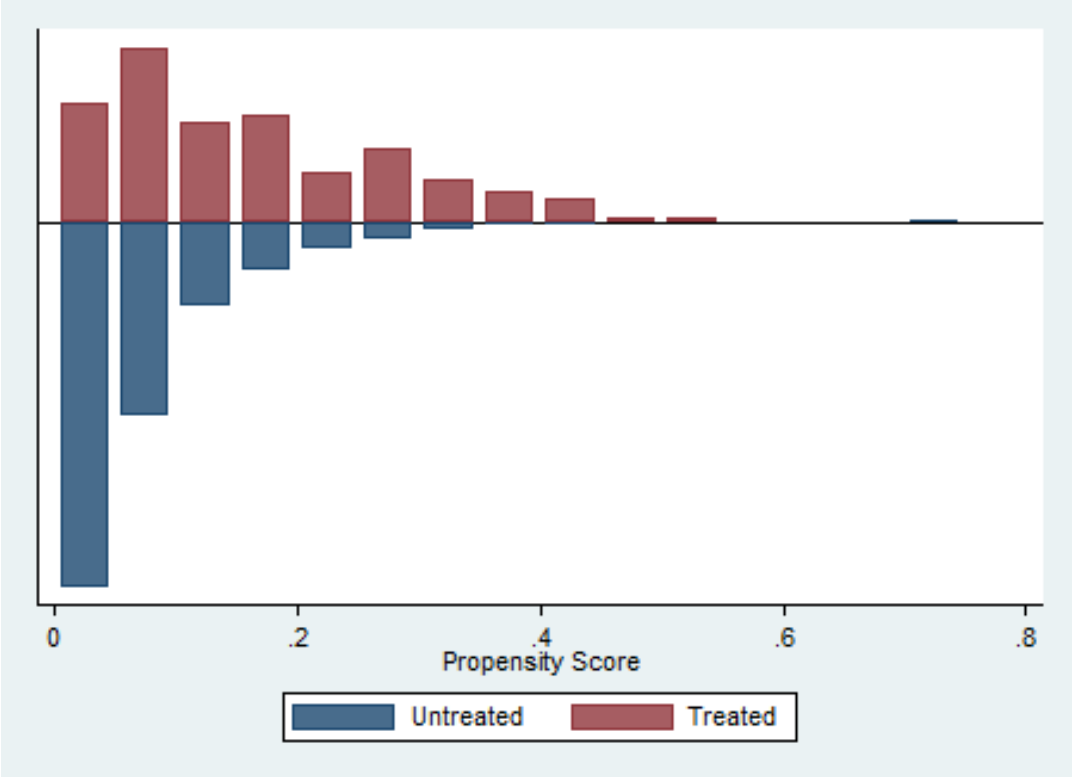
Source: McKenzie (2012b), based on Sankoh et al. (1997) and used in Aker et al. (2011).

B3. Probit results of first-stage propensity score matching procedure

Active FREG membership (Probit)	
Gender hh head *(0-female, 1-male)*	-.4318
	(0.307)
Age hh head	-.0008
	(0.004)
No. of months hh head away from hh	-.0388
	(0.075)
HH head married *(dummy)*	.3552
	(0.301)
Education hh head *(grades completed)*	.0337**
	(0.164)
Non-farm family business hh head *(dummy)*	-.4796**
	(0.208)
Off-farm wage employment hh head *(dummy)*	.0779
	(0.155)
Houseold size	.0323
	(0.031)
Social capital score *(0-1)*	-.0598
	(0.246)
Help-trust score *(0-1)*	-.5149*
	(0.284)
Father of hh head important in Kebele *(dummy)*	.0428
	(0.116)
Walking distance to nearest FTC *(in minutes)*	-.0002
	(0.002)
No. of times talked to DA in the last 12 months	.0079*
	(0.005)
Attended agric. training in the last 12 months *(dummy)*	.6772***
	(0.111)
Basic assets score *(0-4)*	.0274
	(0.067)
Radio owned *(dummy)*	-.2092*
	(0.121)
Food insecurity score *(0-food secure, 12-severly food insecure)*	-.0699***
	(0.021)
Received support from social programs *(dummy)*	.1802
	(0.131)
Total land size in ha	.0857*
	(0.052)
Tropical livestock unit (TLU)	-.0304
	(0.023)
No. of quickwin technologies adopted	.1329**
	(0.055)
Grows main crop *(dummy)*	.1961
	(0.330)
Constant	-1.6775***
	(.500)
Pseudo R-squared	.128
Observations	1,402

Note: Probit results for prediction of active FREG membership in treatment groups, with explanatory variables measured at baseline. 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.

B4. Histogram of estimated propensity score, using nearest-neighbor matching



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.

Ashraf, N., Giné, X. and D. Karlan, 2009. Finding missing markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in Kenya. American Journal of Agricultural Economics 91 (4): 973-990.

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

Bandiera, O. and I. Rasul, 2006. Social networks and technology adoption in Northern Mozambique. The Economic Journal 116: 869-902.

Bandura, A., 1994. Self-efficacy. In: Ramachaudran, V. S. (ed.): Encyclopedia of Human Behavior 4: 71-81. New York: Academic Press.

Bandura, A., 1977. Self-efficacy: Towards a unifying theory of behavioral change. Psychological Review 84 (2): 191-215.

Bandura, A., 1971. Social Learning Theory. New York: General Learning Press.

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.

Bishaw, B., 2001. Deforestation and Land Degradation in the Ethiopian Highlands: A Strategy for Physical Recovery. Western Michigan University: Center for African Development Policy Research (CADPR).

Bloom, H. S., 2006. The Core Analytics of Randomized Experiment in Social Research. MDRC Working Paper. New York: MDRC.

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.

Carter, M. R., Laajaj, R. and D. Yang, 2016. Subsidies, Savings and Sustainable Technology Adoption: Field Experimental Evidence from Mozambique. Working Paper. University of California, Davis/Universidad de los Andes/University of Michigan/NBER/BREAD.

Carter, M. R., Laajaj, R. and D. Yang, 2013. The impact of voucher coupons on the uptake of fertilizer and improved seeds: Evidence from a randomized trial in Mozambique. American Journal of Agricultural Economics 95 (5): 1345-1351.

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.

Conley, T. and C. Udry, 2001. Social learning through networks: The adoption of new agricultural technologies in Ghana. American Journal of Agricultural Economics 83(3): 668-673.

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.

Doss, C. R. and M. L. Morris, 2001. How does gender affect the adoption of agricultural innovations? The case of improved maize technology in Ghana. Agricultural Economics 25: 27-39.

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).

Duflo, E., Kremer, M. and J. Robinson, 2011. Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. American Economic Review 101: 2350-2390.

FAO, 2018. The State of Food Security and Nutrition in the World. Rome: Food and Agriculture Organization of the United Nations (FAO).

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

FAO, 2013. Climate-Smart Agriculture – Sourcebook. Rome: FAO.

Feder, G., Murgai, R. and J. B. Quizon, 2004a. 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.

Feder, G., Murgai, R. and J. B. Quizon, 2004b. Sending farmers back to school: The impact of farmer field schools in Indonesia. Review of Agricultural Economics 26 (1): 45-62.

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.

Giné, X. and D. Yang, 2009. Insurance, credit, and technology adoption: Field experimental evidence from Malawi. Journal of Development Economics (89): 1-11.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M. and C. Toulmin, 2010. Food security: The challenge of feeding 9 billion people. Science 327: 812-818.

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.

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.

Krishnan, P. and M. Patnam, 2013. Neighbors and extension agents: Who matters more for technology adoption? American Journal of Agricultural Economics: 1-20.

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

Matuschke, I. and M. Qaim, 2009. The impact of social networks on hybrid seed adoption in India. Agricultural Economics 40: 493-505.

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.

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.

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.

UN, 2017. World Population Prospects: The 2017 Revision – Key Findings and Advance Tables. Working Paper No. ESA/P/WP/248. New York: United Nations, Department of Economic and Social Affairs, Population Division.

Van Campenhout, B., Vandevelde, S., Walukano, W. and P. van Asten, 2017. Agricultural extension messages using video on portable devices increased knowledge about seed selec-

tion, 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.

Wuepper, D. and J. Sauer, 2016. Explaining the performance of contract farming in Ghana: The role of self-efficacy and social capital. Food Policy (62): 11-27.

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.