

Bonan, Jacopo; Kazianga, Harounan; Mendola, Mariapia

**Working Paper**

## Agricultural Transformation and Farmers' Expectations: Experimental Evidence from Uganda

GLO Discussion Paper, No. 456

**Provided in Cooperation with:**

Global Labor Organization (GLO)

*Suggested Citation:* Bonan, Jacopo; Kazianga, Harounan; Mendola, Mariapia (2020) : Agricultural Transformation and Farmers' Expectations: Experimental Evidence from Uganda, GLO Discussion Paper, No. 456, Global Labor Organization (GLO), Essen

This Version is available at:

<https://hdl.handle.net/10419/213041>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Agricultural Transformation and Farmers' Expectations: Experimental Evidence from Uganda \*

Jacopo Bonan<sup>†</sup> Harounan Kazianga<sup>‡</sup> Mariapia Mendola<sup>§</sup>

January 6, 2020

## Abstract

Why adoption rate of potentially profitable agricultural technologies in Africa remains low is still puzzling. This paper uses a randomized control trial to study Ugandan subsistence smallholders' decisions to adopt cash crops. A unique way of eliciting farmers price and yield expectations allows us to investigate the role of farmers' ex-ante beliefs about crop profitability on adoption decisions. We find that the provision of extension services increases oilseeds adoption by 15%, and farmers who under-estimate oilseeds price at baseline are the most likely to adopt the new crops. The results suggest that changes in expectations drive agricultural technology take-up.

JEL codes: O13, O33, Q14, Q15, Q16

Keywords: Technology Adoption, Commercial Farming, Randomized Controlled Trial, Uganda

---

\*We thank Leah Bevis, Giovanna d'Adda, Jonathan de Quidt, Selim Gulesci, Dean Karlan, Simon Quinn, Diego Ubfal, Kate Vyborny and seminar participants at the UNU–Wider, 2019 NEUDC (Northwestern University), the 2nd International Conference on Applied Development Economics (Lahore School of Economics), the 2nd Scottish Workshop in Development Economics (Heriot-Watt University) and the 1st Workshop on Development Economics (University of Milano Bicocca) for useful comments and suggestions. We are grateful to Connie Masaba and Richard Kabuleta from the Ugandan Ministry of Agriculture, Animal Industry and Fisheries for their superb collaboration with this research. We also thank William Ekere, Franklin Simtowe and Ddungu Stanley for invaluable support throughout the field project and Emanuele Clemente, Federico Maggio and Alessio Romarri for excellent research assistance. We acknowledge financial support from the International Initiative for Impact Evaluation (3ie – Agricultural Innovation Thematic Window). Harounan Kazianga acknowledges also support from the Carson Priority Endowed Professorship. Pre-analysis plan irreversibly registered with 3ie and the Randomized Control Trial registered ex-post in The American Economic Association's registry under RCT ID AEARCTR-0005150. The usual disclaimer applies.

<sup>†</sup>Politecnico di Milano, RFF-CMCC European Institute on Economics and the Environment (EIEE) and Centro Studi L.d'Agliano. Email: jacopo.bonan@polimi.it.

<sup>‡</sup>Oklahoma State University and Centro Studi L.d'Agliano. Email: harounan.kazianga@okstate.edu.

<sup>§</sup>Università di Milano–Bicocca, Centro Studi L.d'Agliano and IZA. Email: mariapia.mendola@unimib.it.

# 1 Introduction

Impediments in commercial agriculture are among the most significant frictions to structural transformation in Africa, where subsistence farming is still predominant. Over the past few decades, agricultural productivity growth has been particularly slow in sub-Saharan Africa as compared to other regions, for which a key explanatory factor is the low adoption rate of new farming technologies and systems (World Bank, 2007; Sunding and Zilberman, 2001; Meiburg and Brandt, 1962). Value chain development, particularly in cash crops, is one potential mean of fostering rural transformation, increasing productivity and enhancing living standards of smallholder households in developing countries (Ashraf et al., 2009; Barrett et al., 2018; Bellemare and Bloem, 2018). Yet, the question as to why some farm households respond to the incentive for agricultural innovation, whereas others do not, is still of great importance to both academics and policy makers. Using data from rural Uganda, this paper investigates subsistence smallholders’ decisions to adopt profitable cash crops (oilseeds) and potentially transition into commercial farming. We exploit the randomized roll-out of a national extension service program to specifically investigate the role of farmers’ *ex-ante* beliefs about crop profitability (i.e. price and yield expectations) in explaining adoption choices.

Since the end of 1990s, the Government of Uganda has been committed to supporting the agricultural sector by heavily investing in a nation-wide vegetable oil extension program supported by both international donors and public-private partnership.<sup>1</sup> The vegetable Oil Development Project (VODP) aimed at increasing farm households’ income by boosting the production of vegetable oil and its by-products for both domestic and regional markets. VODP employed a value chain approach in support of the transition from subsistence to commercial farming by nurturing the commercial relationships between smallholder farmers and processors (millers).

In partnership with the Ugandan Government, we designed and conducted a randomized evaluation of the scale-up phase of VODP (VODP2) in order to evaluate its effects on agricultural productivity and household welfare. The intervention consisted of a package of services, including technical and marketing information about oilseed crops production. We evaluate the one-year effects of VODP2 exploiting the program’s random assignment across administrative units<sup>2</sup> as well

---

<sup>1</sup>The vegetable Oil Development Project (VODP) is a major collaboration between the Ugandan government and the International Fund for Agricultural Development (IFAD). It has remained highly innovative since it introduced a new crop (oil palm) to Uganda along with a new form of public-private partnership. The oil palm component has been concentrated in the southern part of the country and been completely out of our focus. Our analysis pertains to the oilseeds component (i.e. groundnuts, sesame, sunflower, soy bean) that has been concentrated in the North and Eastern area of the country. The total cost of the project is about USD 150 million (2 components), of which USD 52 million is accounted for by the IFAD loan and USD 14.4 million comprises of the Government counterpart. The total number of beneficiaries is nearly 139,000 households. Section provides more details on the program.

<sup>2</sup>The roll-out is across sub-counties, which are at an intermediate administrative level (between districts and

as unique data on farmers’ expectations about oilseeds profitability, i.e. individual beliefs on both expected prices and yields, collected before and after the program.

The approach of eliciting *ex-ante* returns to technology adoption allows us to shed light on the determinants of risky choices, since it is the returns perceived by individuals at the time of their adoption decision that matter for actual choices which we only observe *ex-post* (Manski, 2004; Attanasio, 2009). Indeed, while these beliefs are not necessarily correct, they do play a central role in understanding investment choices. Systematic mis-perceptions about profitability, akin to knowledge and information imperfections, can prevent technology adoption, thereby justifying information-centered interventions.

This paper is related to an extensive literature on the low uptake of new technologies by small-holder farmers in developing countries. Both supply-side (markets and exposure) and demand-side (risk aversion and behavioral biases) constraints have been offered as explanations of low technology adoption (Kondylis et al., 2017; Ben Yishay et al., 2019).<sup>3</sup> Several observational and experimental studies have explored various ways of mitigating obstacles to adoption, such as the paucity of access to product and financial markets and insufficient public investment in roads and other infrastructures (Ambler et al., 2018; Karlan et al., 2014; Stifel and Minten, 2008). More recently, other potentially interlinked explanations have been postulated, including lack of knowledge, time-inconsistent preferences and incomplete learning (Ashraf et al., 2009; Duflo et al., 2011; Hanna et al., 2014).

The role information plays in new technologies’ adoption decisions puts the onus on the performance of extension services, the transfer of knowledge and information failures in the process of learning for adopting (e.g. Feder et al., 1985, 1987; Anderson and Feder, 2008; Waddington et al., 2012; Bonan and Pagani, 2018; Deutschmann et al., 2019).<sup>4</sup> However, technology adoption represents a gamble on precisely which production function will be realized (e.g. Ashour et al., 2019; Bold et al., 2017). Thin and poorly integrated agricultural markets in developing contexts, in conjunction with highly heterogeneous conditions, generate large variations in crop yields and prices of agricultural products (Fafchamps, 1992; Barrett and Dorosh, 1996; Minten and Kyle, 1999; Burke et al., 2018). Hence, farmers use noisy information to form their *ex-ante* expectations about the production function (i.e. forward looking decision making) when making adoption decision.

---

villages) with average 20,000 population.

<sup>3</sup>See Magruder (2018) and de Janvry et al. (2017) for a broader review.

<sup>4</sup>Existing evidence on the impact of extension services shows mixed results, and several studies have faced measurement issues and other challenges in comparing results across programs (Anderson and Feder, 2008; Aker, 2010; Maffioli et al., 2013; Maertens et al., 2018). Even when extension encourages farmers to adopt new practices, the final impact in terms of the measurable increases in yields or profits is rarely detected (Beaman et al., 2013). More recently, other studies have documented the efficacy of delivering market (price) information, possibly through new media devices (Jensen, 2007; Cole and Fernando, 2018; Fafchamps and Minten, 2012; Aker, 2010).

Their experience with the performance of a new technology on their land may be limited and typically, they do not directly observe technology under the full set of states of the world. Even in case of familiar technologies with relatively high returns, learning opportunities may be restricted by substantial heterogeneity across farmers (Suri, 2011).

In the existing literature, evidence on the relevance of heterogeneity in returns to farm technology is rather mixed (e.g. Barrett et al., 2018), and generally comes from indirect methods which aim to overcome the problem of self-selection (e.g. Beaman et al., 2014; Jack, 2013). In this paper, we take a more direct approach in investigating the extent to which heterogeneity in farmers' expected returns determines the decision to adopt by using *ex-ante* (i.e. prior to the adoption choice) farmer level data on beliefs about future returns as well as their change after the information treatment.

Technology adoption and their accompanying potential returns are determined by farmers' preferences, access to resources/inputs as well as other factors that may be determined at a more aggregated level (e.g. the village or the district). Even after controlling for observable, unobservable factors may bias the estimated relationship between technology adoption and realized returns, usually observed *ex-post* (i.e. after actual investment choices). The current study surmounts this difficulty by using two orthogonal sources of variation, i.e. a randomized controlled trial of extension services and *ex-ante* expectations about the returns to adoption. Importantly, while considering farmers' beliefs about expected crop profitability, we distinguish between price and yield expectations, thus ascertaining the relative importance of technical and market sources of (mis)perceptions that are relevant to farmers' decision making process related to adoption. While the role of expectations in human capital investment decisions (e.g. health, education, migration) has been recently explored in the literature (Jensen, 2010; Attanasio and Kaufmann, 2014; Wiswall and Zafar, 2015; Delavande and Zafar, 2018; McKenzie et al., 2013; Delavande and Kohler, 2016), our analysis provides what is, to the best of our knowledge, the first characterization of heterogeneous farmer expectations in agricultural innovation and technology adoption.

We estimate that the program increased adoption of oilseeds by 3.7 percentage points and the share of land allocated to oilseeds by 0.6 percentage points, which correspond to an increase of about 15% and 17.5% relative to the control group at endline. We find a significantly positive impact on a number of technical outcomes including use of fertilizers and improved seeds, non-family labor, access to markets and marketing strategies (e.g. bulk sales), cash crop production and productivity. Yet, we report modest effects on farm revenues and household welfare outcomes, which is consistent with results from recent studies that did not detect a significant effect on farm profits (e.g. Beaman et al., 2013) and household economic well-being in the short-run (e.g. Carter et al., 2014).

Importantly, we find significant heterogeneous effects on cash crop adoption. Farmers who under-

estimate oilseed price at the baseline (low expected price) are more likely to grow the new crop. In particular, farmers whose expected price at baseline is lower than the actual market price (positive price wedge) are more likely to adopt after being treated. In contrast, heterogeneity in yield expectations did not have any detectable effect. Finally, we find that the program induces farmers to significantly change and correct their beliefs about profitability *ex-post*, particularly by reducing the wedge between expected price and actual retail prices. This holds especially for those systematically underestimating the actual price at baseline, pointing to farmers revising their beliefs in a logical way.<sup>5</sup>

The exercise on heterogeneity leverages random variation in both technical and market information and on pre-determined, non-randomized, variation in expectations. This raises the concern that unobservables correlated with both expectations and other determinants of oilseed adoption might drive our findings. We tackle this issue by investigating observable possible sources of heterogeneity in oilseed adoption, such as familiarity with oilseed and wealth, and our results are robust to controlling for them simultaneously with our measures of expectations.

Our findings are in line with recent studies which show that farmers in developing countries respond to innovation incentives when there are positive expected returns, even though the prevailing uncertainty about profitability plays a crucial role in both take-up decisions and subsequent investment (Barrett et al., 2018; Oliva et al., 2019). In a recent experimental study in Benin, Arouna et al. (2019) establish the importance of price uncertainty in technology adoption. The study randomizes different contract farming arrangements and the one that only includes an agreement on price has the largest effect on smallholders' performance, as large as other types of contracts with different attributes in order to eliminate technical constraints. Similarly, an extensive body of literature has documented that, if the technology is appropriate and expected profits are sufficiently high, farmers can overcome obstacles to adoption, including bad institutions and poor roads<sup>6</sup>.

Our results show that increased access to information through extension services has a positive impact on cash crop adoption and technical changes among smallholder farmers in Uganda. In particular, low information (access) is mitigated by the provision of extension services, which enable farmers to revise and correct their beliefs about crop profitability (in a logical way). We show that farmers' expectations on prices, but not on yield, are the main mechanisms underlying the adoption switch, pointing to the key role of price uncertainty and market information in smallholders' technology upgrade.

---

<sup>5</sup>Heterogeneous effects of profitability of expectations on oilseed adoption was not explicitly mentioned in the pre-analysis plan but such dimension is listed among primary outcomes, in conjunction with adoption, improved farming techniques and market linkages, in our theory of change.

<sup>6</sup>See discussion by Udry and Gollin <https://assets.publishing.service.gov.uk/media/5c44c3d3ed915d3890c53f9f/Raising-agricultural-productivity-in-sub-Saharan-Africa.pdf>

Ultimately though, increased information leads to enhanced uptake but not performance, at least in terms of short-term household revenues and income. One possible explanation is that increasing revenues and income requires that farmers be properly integrated into the downstream part of the value chain and doing so would take more than a year. Nevertheless, the lack of significant effects on revenues and income may represent a threat to the 'follow-through stages' in oilseed investment and disadoption may occur if expected returns to technology adoption are not realized over several years.

The rest of the paper is organized as follows. In Section 2, we describe the context and the details of the VODP intervention along with the experimental design we developed. Section 3 summarizes our panel survey data, while section 4 presents the empirical strategy and experimental results. Thereafter, Section 5 presents findings related to the mechanisms underlying the results. Section 6 discusses the empirical results and concludes.

## 2 Experimental Design and Data

### 2.1 Program description

Agriculture is one of the mainstays of Uganda's economy, accounting for 22 per cent of gross domestic product (GDP) and engaging two-thirds of the economically active population (UBOS 2010, 2013). The vegetable Oil Development Project (VODP) is a flagship governmental development project that was commenced in 2000 in collaboration with the International Fund for Agricultural Development (IFAD). One key feature of this program is the strong leadership of the Government, through the Ministry of Agriculture, Animal Industry and Fisheries (MAAIF), which created a conducive environment for the first large public-private partnerships in agribusiness in Uganda.<sup>7</sup>

The overall goal of VODP is to increase household income of smallholders by revitalizing and increasing domestic vegetable oil production, in partnership with value chain stakeholders (i.e. millers and traders, input suppliers and cooperatives and other producer organizations and services). To that end, this project helps farmers increase their production of crushing seeds and establishes commercial relations to connect them directly to processors. VODP is highly relevant to government policy, within the Plan for Modernization of Agriculture, as a source of growth, poverty reduction,

---

<sup>7</sup>The project is funded by the Government of Uganda (GoU), the International Fund for Agricultural Development (IFAD) as well as private actors with contribution from major stakeholders following public-private partnership arrangements. VODP was approved by IFAD's Executive Board and by the Parliament of Uganda. Originally, the total cost of this project was supposed to be USD60 million, consisting of an IFAD loan of USD20 million, USD33.1 million of co-financing from the private sector partner, USD3.8 million from the Government of Uganda and USD3.1 million from beneficiaries. However, due to an increase in the scale of the project, the private investor and the Government raised their contributions to USD120 million and USD12 million, respectively, thus raising the total cost to about USD156 million.

as a mean of fostering partnerships with the private sector for promoting import substitution and export diversification.

VODP was developed in two phases: the first one (VODP1) ran for eleven years (until the end of 2011) and was followed by a scale-up extension service program (VODP2), which was expected to begin in new areas in 2014.<sup>8</sup> The oilseeds component of VODP2 targets groundnut, sunflower, soybean and sesame, and aimed at reaching 136,000 smallholder households around four hubs (Lira, Eastern Uganda, Gulu and West Nile) covering 43 districts in Uganda (see Figure 1).

VODP2 extension model entails the development of a competitive oilseed value chain by enhancing smallholders' production and inclusion. Extension services are supplied by private service providers and include three components: group formation and strengthening, technical services for oilseed production and productivity, information about commercial/marketing strategies, market information, access and linkages. One significant technical component is the farmer learning platform, which is a form of applied research and technology transfer system on the latest notified/released varieties along with full package of practices on selected farmers' fields. The objective is to demonstrate the potentiality of the underlying technologies to both attending farmers and neighboring farmers/agencies.<sup>9</sup>

Project activities are targeted at farmer groups. The latter may be already existing and active groups or created for the purpose of the project. On average, groups comprise of 15 to 20 male and members, and have a leader. Farmer groups are eligible to receive VODP2 project and to participate our study when: i. They have been already formed for other previous projects and have some degree of formalization (formal constitution, registration at the level of sub-county or district city office); ii. They are interested in oilseed production; iii. They have available land in order to implement the learning platforms; iv. They are not currently benefiting from other development projects.

Due to both resource constraints in reaching all areas at the same time and for the purposes of impact evaluation, the choice was made to gradually deploy VODP2 activities to the targeted beneficiaries. Hence, in partnership with the Ugandan Government, we designed VODP2 with a phase-in structure, which allowed us to implement a randomized controlled trial.

---

<sup>8</sup>The actual implementation of the program was delayed and began from 2016.

<sup>9</sup>The primary components of each block are: 1. technical services for increased oilseed production/productivity: technology demonstration and farmer training on best agronomic practices; conservation agriculture; integrated soil fertility management; integrated pest and disease management; post-harvest handling and storage; farm power and mechanization such as the use of ox-ploughs and tractors; complementary enterprises such as apiary and livestock 2. market linkages and access: bulking for produce and inputs; market information gathering and market intelligence; negotiation and negotiation skills development; building commercial relationships with the market such as linkages with seed companies, input dealers, oilseed millers, and financial institutions 3. group formation and strengthening: mobilization and sensitization; farmer group needs assessment; training in group dynamics and management; action planning; farmer group institutionalization through registration.



## 2.2 Sampling

The evaluation uses a phased roll-out of the VODP2 and employs a cluster-randomized block design, where the blocks consist of sub-counties<sup>10</sup> and clusters consist of farmer groups. The sampling follows a multi-stage process for selecting sub-counties, farmer groups and farmers in turn. We first randomly select a sample of 86 sub-counties in 15 districts to be assigned to the treatment and control group. The selection is made from a list of 130 sub-counties deemed eligible for expansion of VODP2 in the Northern and Eastern regions, provided by VODP2 management unit (MAAIF). The selection is based on orographic and soil characteristics of the area in relation to the development of oilseed production. In the final distribution of sub-counties, 23 and 63 are in the Northern and Eastern regions, respectively.

The second level of sampling pertains to the selection of farmer groups, within each sub-county. At this stage, we sample 8 farmer groups per sub-county, for a total of 688 groups. A list of farmer groups that meets the eligibility criteria is obtained from the service providers contracted to deliver VODP2, in collaboration with community development offices.<sup>11</sup> From the list of eligible farmer groups, 8 groups per sub-county are randomly selected.<sup>12</sup>

The third level of sampling pertains to the selection of farmers within farmer groups. At the time of the household survey, we sample 4 farmers per group, for a total of 2752 farmers.<sup>13</sup> After the selection of a farmer group, a meeting with the leader or a representative of the group is set up by our survey team. During the meeting, the group level questionnaire and the household level questionnaire are administered to the group leader. Accordingly, the list of other group participants is obtained, and 3 members are randomly selected in order to be interviewed. A total of 688 questionnaires are administered to group leaders or representatives and 2064 to other members; overall our sample should be made by 2752 farmers. Net of missing observations, our

---

<sup>10</sup>Uganda regions are divided into 112 districts. Each district is further divided into counties and municipalities, and each county is in turn divided into sub-counties. According to the 2016 General Elections Statistics, the total number of sub-counties is 1403 (with on average 12,5 sub-counties per district). The average sub-county population is 26,111, with a lot of variation across them. Finally, in the entire country, 5,842 villages are counted (41 village per sub-county on average).

<sup>11</sup>At that time, our sole focus was on the already existing groups.

<sup>12</sup>Even though the instructions for the random selection of groups have been made clear to all stakeholders, we could not directly oversee the process. Therefore, we cannot guarantee that this was always strictly implemented. However, this should not be a threat to our identification of impacts because the random assignment to treatment and control group was performed after the identification of farmer groups and the baseline survey. Yet, if, for instance, the selection process was biased towards most promising groups, as long as that was the case in all sub-counties, it would cast doubts on the external validity of our results, but not on the internal validity.

<sup>13</sup>Decisions over sample sizes at all levels are guided by ex-ante power calculations. In particular, we base our calculations on groundnut yield from Uganda reported by the World Bank Living Standard Measurements Surveys (LSMS). Groundnut was chosen because it is one of the main crops that is targeted by the project and for which yield information is easily accessible. Sample size was selected in order to detect standardized effect of 0.20 at the significance level of five percent with a statistical power of 0.80.

working sample comprises of 2733 valid farmer level questionnaires.<sup>14</sup>

## 2.3 Experimental design and data

For the purpose of impact evaluation, we exploit the roll-out structure of the VODP2 to implement a randomized controlled trial. We focus on two oilseeds hubs in Uganda (Mbale-Jinja and Gulu region) in the Eastern and Northern regions, respectively. They are encompassed by three private service providers contracted by VODP2 management unit (MAAIF) to deliver the oilseed extension program following the same terms of references. The research design was developed jointly with VODP2 and service providers. The program is randomized at the sub-county level, which makes it easy to reconcile program implementation constraints and research requirements. The phase-in design involves a sample of eligible sub-counties that are randomly assigned to receive the treatment in different points in time. In particular, one random group of sub-counties (treated group) is assigned to receive the entire package of project interventions during 2017/early 2018. The second group (control group) was expected to receive the treatment in late 2018/early 2019 after the endline survey administration.

We assigned the treatment, i.e. the moment when sub-counties receive the intervention, after the baseline, stratifying at the level of district while using a random number generator. The list was made available to VODP2 and service providers, but was not publicly disseminated. This, in turn, allows us to rule out any type of selection arising from the (in)correct application of eligibility criteria for group inclusion and sampling issues. In case there is a mistake in group selection, this would not jeopardize the internal validity of our exercise, as randomization occurred at the later stage. Figure 1 illustrates a map of the study areas and of the random assignment of sub-counties.

The surveys were fielded at both farmer and group level. The latter includes information on the characteristics of the group (such as size, composition, scope, functioning and activities regularly carried out). Farmer-level surveys gather information on general farm-household characteristics, agricultural production (during the two previous seasons), skills and performance along with several socio-economic variables. At the endline, we add questions to address the extension service activities performed/received by both farmers and groups. Baseline data were collected between June and September 2016 from all farmer groups, before any activity started in any sub-county that was included in this study. The endline data were collected between July and October 2018, using the same survey instrument as the baseline. Figure 2 illustrates the study timeline.

As is the case with any longitudinal survey, attrition was unavoidable. Attrition rate in this survey was found to be relatively modest as of 7%. In absolute terms, of the 2733 farmers who were

---

<sup>14</sup>The reasons for the impossibility to use nineteen questionnaires are: severe data incompleteness and data loss. We do not observe any systematic pattern in the geographic distribution of such cases.

surveyed at baseline, 189 were not resurveyed at endline. In Appendix A we show that natural attrition is not differential across treatment and control groups. In addition, it is not significantly predicted by any relevant covariates at baseline. We conclude that attrition was random, and hence does not threaten the internal and external validity of our analysis.

### 3 Descriptive Statistics and Balance Tests

#### 3.1 Respondent and household characteristics

In this sub-section, we use the baseline data to test the balance of the randomization experiment. We present the mean of the specific variable measured at the baseline for the control group as well as its difference from the treated group.<sup>15</sup> Detailed descriptions on variables construction are provided in Appendix B.

We summarize household level variables in Table 1. Column (1) reports the mean and standard deviation for the control group at the baseline and column (2) tests the initial randomization balance between the treatment and control group. Panel A reports respondent characteristics, while panel B shows household level general outcomes (some of these variables also find inclusion in the vector of controls in the regression analysis described below). Respondents are mostly household head (61% of the cases) and about half of them comprise of male (52%). 73% of respondents who can read and write. Less than half (47%) have completed primary education, 37% and 5% have attained secondary and higher education, respectively, while 11% have no formal education at all. Households cultivate, on average, 6.6 acres of land organized in about 2.3 plots and dedicate about 233 days to farming activities every year. Seasonal revenues from crop sale are about 1,330,000 Ush while the typical household monthly income is around 230,000 Ush<sup>16</sup>. We also consider a set of wealth indicators (including livestock and household assets), which we aggregate in a wealth index using the first component obtained with Principal Component Analysis (PCA) (Filmer and Pritchett, 2001). Column (2) illustrates that the randomization was largely successful since we do not detect any significant imbalances between treatment and control groups for both the individual and household level characteristics.

#### 3.2 Oilseeds-specific outcomes

In Table 2, we show the crop specific balance tests. The layout is similar to that of Table 1, with the exception that four oilseed crops are now shown in each pair of columns. In general,

---

<sup>15</sup>The exercise is conducted by regressing each baseline variable on the treatment dummy, with robust standard errors clustered at the sub-county level.

<sup>16</sup>These figures correspond to 357 USD and 62 USD, respectively, at 2018 market exchange rate.

variables aggregate information coming from two seasons for each survey wave. More details on the construction of variables are provided in the Appendix B.

Panel A of Table 2 summarizes the baseline adoption of oilseed crops at baseline at the extensive (adoption rate) as well as at the intensive (share of land cultivated with oilseeds) margins. Groundnuts are the most adopted crop, and are cultivated by 43.7% of the households on about 10.9% of farmland. Soybean and sesame are each farmed by about 20% of our households sample on approximately 3.1% and 4% of farmland, respectively. Sunflower, the least grown crop, is cultivated by 10.8% farmers on 2.6% of farmland.

The utilization of modern inputs on oilseed crops, summarized in panel B, varies substantially across crops. The figures exhibit two key features. First, households consistently use more modern inputs on soybeans and groundnuts, and relatively little modern inputs on sunflower and sesame. Second, input usage is not positively correlated with adoption. As an example, the adoption of soybean and sesame is around the same at the extensive and the intensive margins, but households applied 367 kilograms of fertilizers and 370 kilograms of pesticides on soybean as opposed to 6 kilograms of fertilizers and pesticides on sesame. Overall, the pattern is suggestive of complementarity amongst modern inputs rather than sustainability, or at least as per the households' perception.

Panel C provides an overview of labor contribution from various sources and its allocation across the four oilseed crops. Total labor applied varies from 12.4 days for groundnuts to 3.5 days for sunflower. The head and his/her spouse supply roughly an equal amount of labor on each crop. The labor market is thin, with less than a day of outside labor used oilseed crops, with the exception of groundnuts on which about 1.7 days of outside labor is used.

In panel D, we tabulate market access indicators, all of which are dummy variables comprising of each specific oilseed crop, any sale of farm produce, purchase of seeds, bulk and group sale of farm harvest, contact with value chain actors and sale to market actors. We find that 28% of households cultivated groundnuts reported selling some of their harvest, the highest of any oilseed crops followed by soybean (14%), sesame (12%) and sunflower (8%). We observe a consistent pattern for the remaining market access indicators, with groundnuts and soybean being more integrated in the market as compared to sesame and sunflower.

We summarize production and productivity outcomes in panel E. Groundnuts appear to be the most productive crops with 170 kg per acre. Yields per acre average less than 100 kg for soybean, sunflower and sesame. We observe the same pattern for total harvest (both in quantity and in monetary value) and sale revenues, which indicates that groundnut is the most significant crop.

Taken together, the figures in Tables 1 and 2 indicate that the randomization exercise succeeded in generating treatment and control groups with similar characteristics at baseline. Looking at the t-tests, we fail to reject the null hypothesis that the difference between treatment and control is zero

in only 4 cases (or 3.1% of all the difference tests), precisely for ‘use of and expense for pesticides for groundnuts in Panel B of Table 2, and “harvest in bulk/group” for soybean in Panel E. We impute this to chance. Moreover, using a test of joint orthogonality, we are unable to reject (at any conventional level) the null hypothesis that the average characteristics of farmers and their plots are unrelated to assignment to treatment<sup>17</sup>.

### 3.3 Expectations about oilseed profitability

The individual level survey includes original data on farmers’ expectations on oilseed profitability. Importantly, we consider both price and yield expectations in order to distinguish the market- *vs* technical-related beliefs. We ask respondents about their beliefs pertaining to *own* expected price and yield (per unit) at the end of the season if they were to grow each specific oilseed. We further collect data on their beliefs about the average price and yield (per unit) faced by the *average farmer* in the rest of the population.<sup>18</sup>

In order to measure the degree of bias in *ex-ante* beliefs on oilseed price and yield, we calculate the wedge between farmer’s belief and the actual price and yield of each specific oilseed. For the former, we use season and crop-specific retail prices, measured at the level of districts, provided by Info Trade-AGMIS<sup>19</sup> For yields, we use the data on the average yields from the latest Uganda 2015/2016 Living Standard Measurement Surveys (LSMS) data, by focusing on the districts that are included in our study. Wedges are expressed as percentage deviations from actual values, i.e. they are built as the difference between actual and expected values, divided by the former.

Average expected yields, prices and their respective wedge for both treatment and control group can be found in panel F of Table 2 and appear balanced across the treatment and control group. At baseline, farmers were generally misinformed about oilseed yields and prices. Specifically, respondents were over-estimating yields and under-estimating prices (i.e. average yield wedge is negative while prices wedge is positive). Expected yields per acre were over 700 kg for sunflower and groundnuts, 531 kg for soybeans and 393 kg for sesame, revealing enormous differences with

<sup>17</sup>Results are not shown but available upon request.

<sup>18</sup>We collected farmers’ expectations about oilseed profitability from the following questions, which are asked for each oilseed (i.e. groundnut, sunflower, soy bean, sesame): (i) *Consider the hypothetical situation where you grow [CROP]. Look ahead after the harvest, how much do you think the end-of-the-season PRICE (US\$/Kg) of [CROP] will be for you?* (ii) *Consider the hypothetical situation where you grow [CROP]. Look ahead after the harvest, how much do you think the average YIELD (Kg/Acre) of [CROP] will for you?* In addition to the questions above, farmers were asked the following questions: *“Consider a typical farmer growing [CROP]. How much do you think the end-of-season price (US\$/Kg) of [CROP] can be?”* and *“Consider a typical farmer growing [CROP]. How much do you think the end-of-season yield (Kg/Acre) of [CROP] can be?”*

<sup>19</sup>The price data are taken from <https://agmis.infotradeuganda.com> We aggregate monthly prices into annual prices. Price data were available for 7 out of 15 districts. For the remaining 8 districts, we carry out the extrapolation by using the average price data from the nearest three districts.

the measured yields, as shown in panel E. The over-estimation is captured by the wedge values. For instance, groundnuts’ expected yield as reported by respondent was more than 200% higher than actual yield. Expected prices were lower than actual ones, except for sesame. Moreover, the relative errors in price expectation averaged from 31.4% for sesame to 41.5% for groundnuts.

Collecting data on individual beliefs about expected returns of future hypothetical choices poses some challenges, particularly in developing country settings, where respondents may have limited formal schooling or training. Yet, there is a growing body of research which suggests that such an endeavor is indeed possible if enough care is given to the design of the questionnaires, and this is even more so if the focus is on the first-moment of the expected distribution of returns (Delavande et al., 2011; Attanasio, 2009). We employ common protocols, trained enumerators and run accurate survey piloting (in the same areas where we run the actual survey).

Despite these efforts, the proportion of missing values for profitability expectation variables is non-trivial. We formally investigate the determinants of missing values and its consequences on results in Appendix A. First, we find that there is no correlation between the missing values in price and yield expectations at the baseline. Second, as one would expect, response rate significantly increases in familiarity with oilseeds, education levels and wealth. This has implications for the external validity of the results, including these dimensions, whose application should be confined to the most progressive and relatively well-off part of the study sample.

Given the difficulties in eliciting subjective expectations, we further check both internal and external validity of our expectations data. To start with, since expectations are based on current information, which can be a mixture of public and private information, we collect data on both ‘self-beliefs’ and ‘population beliefs’, as mentioned above (see footnote 18). Hence, we compare the two types of individuals’ beliefs in order to check whether our sample farmers report self-beliefs that are systematically different from expectations related to the ‘average farmer.’ Moreover, since extension services provide a type of public information (related to the average farmer), it is important to determine whether population beliefs are significantly correlated to beliefs about self-returns (if not, that may be a reason for non-relevance of the extension program). Differently said, farmers may be affected by a public information program only if their self-beliefs are connected to their ‘population-beliefs’ (Wiswall and Zafar, 2015). In columns 1 and 2 of Table 4, we find that our own expectations are positively and significantly associated with population beliefs.<sup>20</sup>

In order to further assess the internal validity of the information on subjective beliefs, we test the extent to which subjective expectations co-vary with observed respondents’ characteristics (Attana-

---

<sup>20</sup>We repeat the exercise in Table 4 after including the usual set of controls used throughout the paper and find qualitatively similar results (the coefficient in column one slightly decreases with p-value=0.12. Results are not shown but are available upon request.

sio and di Maro, 2008; Attanasio, 2009). We proceed by regressing baseline subjective expectations on prices and yield on the characteristics of observable baseline farmers. The estimation results are reported in Table 3. In column 1 and 4, we report the coefficients of the main set of covariates used throughout the analysis. In columns 2 and 5, we add baseline adoption. Meanwhile we report the coefficients attached to the KKK indexes for all families of outcomes at the baseline columns 3 and 6. We find in columns 1-3 that none of the covariates, when taken individually, is correlated with price expectations. Significant associations however, are found between yield expectations and a number of variables in columns 4-6. In particular, higher expected yields are found to be positively and significantly associated with being the head of the household, female, more educated, wealthier (both in terms of assets and land) and with deeper familiarity with oilseeds, intended as higher probability of adoption, technical experience (higher productivity) as well as better linkages to the market. These estimation results provide useful insights on the manner in which farmers tend to form their beliefs on prices and yields. At baseline, farmers have similar price information that they use to form their price expectations. To the extent that these smallholders are all price takers, it is not surprising that price expectations are not correlated with any of the observable farmers characteristics. Unlike prices (that are determined by the market), yields are presumably farmer and even farmer- plot specific. For example, a household head who could allocate himself the best household land would expect higher yields than another household member who is allocated marginal land.

Checking external validity of the subjective expectations data usually involves more challenges since data requirements for such an exercise are very demanding. Indeed, matching expectations data with actual realizations would pave the way for external validation, but time-series expectation data are rather rare and unique (see Wiswall and Zafar, 2015). We collect such data before and after the controlled trial, so that we can assess the extent to which subjective expectations predict actual outcomes at the baseline in the first test. Table 4 shows the results of a model where actual yields and prices<sup>21</sup> are regressed on yields and price own expectations. We find that yield expectations are positively and strongly significantly related to the actual realizations, measured both through our survey and external data sources. Price expectations are also positively associated with actual prices, although they are less precisely estimated.<sup>22</sup>

---

<sup>21</sup>Actual yields at the farmer level taken from the survey are used in column 3, while actual yields from LSMS are used in column 4. Actual prices, in column 5, are taken from Trade-AGMIS.

<sup>22</sup>We repeat the exercise after including the usual set of controls used throughout the paper and find qualitatively similar results. Results are not shown but are available upon request.

### 3.4 Project take-up

Table 5 summarizes take-up rates of the project activities. The figures therein indicate that 71.5% and 24% of farmers in the treatment and control groups, respectively, have been actually reached by VODP2. This partial compliance appears to be applicable to most of the variables elucidating the contents that are included in the extension package, i.e. field training on specific oilseed crop, technical advice and market activities. It is possible that other extension projects were active in both treated and control areas during the study period and farmers were unable to reliably identify the program they received. This may limit our confidence in attributing project activities received to VODP2 exclusively. Remarkably, the treated group appears to be significantly more exposed than the control group along all variables, with the exception of the distribution of free inputs by service providers.<sup>23</sup> However, the latter aspect strengthens the impact of the program upon adoption. If the project was providing free seeds and inputs to the beneficiaries, the following adoption would result as a mechanical effect that would cloud farmers’ true intention.

In wake of the partial compliance, in our empirical analysis, we report both the intent to treat (ITT) effects and the local average treatment effects (LATE). For the latter, we use the variable “VODP2 program take-up” in the models for general farmer level outcomes, and the set of specific training dummies in oilseed-specific models, as explained in the following section.

## 4 Empirical Strategy and Results

### 4.1 Specifications and identification

We start the impact analysis by focusing on the intent-to-treat (ITT) effect, which measures the average effect of being randomly assigned to VODP2, regardless of actual take-up. Given that the extension service program specifically focuses on four oilseeds (groundnut, sesame, soy bean, sunflower), we run a pooled regression with crop fixed-effects as follows:

$$y_{icjd,t=1} = \alpha_0 + \beta_1 \text{Treat}_j + \gamma X_{icjd,t=0} + \delta y_{icjd,t=0} + \omega_c + \mu_d + \varepsilon_{ijd} \quad (1)$$

where  $y_{icjd,t=1}$  denotes farmer’s  $i$  outcome of interest for oilseed  $c$  in sub-county  $j$  in district  $d$  at the endline;  $\text{Treat}_j$  is an indicator variable that takes the value one if the sub-county was assigned to the treatment group, and zero otherwise,  $\omega_c$  are crop fixed-effects and  $\mu_d$  are district fixed effects, where districts are the stratification variable.  $y_{icjd,t=0}$  represents the outcome measured at the baseline, whereas  $X_{icjd,t=0}$  is a set of individual and household level controls, including gender,

---

<sup>23</sup>VODP2 did not envisage any input distribution, with the exception of the inputs that are deemed necessary to set up the learning platform.



education, land size, agronomic skills, and household wealth index, all of which are measured at the baseline. The parameter  $\beta_1$  identifies the ITT effects. Standard errors are clustered at the level of sub-county, which is our unit of randomization. We explore heterogeneity effects across different dimensions measured at the baseline by adding the main and interaction terms to the equation above.

We further estimate the crop-specific LATE using the following system of two equations, whereby the treatment assignment is used as an instrumental variable for receiving VODP2.

$$y_{icjd,t=1} = \lambda_0 + \eta_1 \text{VODP2}_{icjd} + \gamma X_{ij,d,t=0} + \delta y_{icjd,t=0} + \omega_c + \mu_d + \varepsilon_{icjd}^a \quad (2a)$$

$$\text{VODP2}_{icjd} = \pi_0 + \pi_1 \text{Treat}_j + \pi X_{ij,d,t=0} + \pi_2 y_{icjd,t=0} + \omega_c + \mu_d + \varepsilon_{icjd}^b \quad (2b)$$

where  $\text{VODP2}_{icjd}$  indicates whether farmer  $i$  in sub-county  $j$  in district  $d$  self-report to receive oilseed-specific extension services  $c$  within the VODP2 project. All the other variables are defined as above. We use the system of equations (2a) and (2b) to estimate the LATE, using  $\text{Treat}_j$  serving as an instrument for the take-up variable in  $\text{VODP2}_{icjd}$ . We report the F test of excluded instruments, as a test for weak instrument.

Identification of the LATE rests on some assumptions (Angrist et al., 1996; Imbens, 2014; Imbens and Angrist, 1994). The first one is being allocated to the treatment increases the probability to take-up. We test the first stage condition by estimating equation (2b) in Table C.1, in Appendix C. Second, the assignment in VODP2 must be random. The randomized control design along with Section 3.1 above provides evidence to support the assumption that the assignment to treatment is random.

Finally, for the set of outcomes that are not oilseed-specific, we estimate farmer level regressions, similar to (1) for ITT:

$$Y_{ij,d,t=1} = \alpha_0 + \beta_1 \text{Treat}_j + \gamma_0 X_{ij,d,t=0} + \delta_0 Y_{ij,d,t=0} + u_d + \epsilon_{ij,d} \quad (3)$$

and similar to (2) for LATE:

$$y_{ij,d,t=1} = \lambda_0 + \eta_1 \text{VODP2}_{ij,d} + \gamma X_{ij,d,t=0} + \delta y_{ij,d,t=0} + \mu_d + \varepsilon_{ij,d}^a \quad (4a)$$

$$\text{VODP2}_{ij,d} = \pi_0 + \pi_1 \text{Treat}_j + \pi X_{ij,d,t=0} + \pi_2 y_{ij,d,t=0} + \mu_d + \varepsilon_{ij,d}^b \quad (4b)$$

where  $\text{VODP2}_{ij,d}$  indicates whether or not farmer self-report to take up any activities within the VODP2 project. All the other variables are defined in the following manner.

We account for a number of outcomes using multiple hypothesis testing in two ways. First we create a family-specific index for all outcomes in the family. Such index, proposed by Kling

et al. (2007), henceforth KLK index, is constructed as the equally weighted average of z -scores of its components, with the sign of each measure oriented so that more beneficial outcomes have higher scores.<sup>24</sup> Second, we test the Family-wise Error Rate (FWER) across indexes by adjusting their respective p-values following the Hochberg step down procedure (Hochberg, 1988). These are reported under the standard coefficients' p-value that refer to the KLK indexes.

As robustness checks, we estimate the following alternative specifications. First, we re-estimate the ITT and the LATE dropping all control variables with the exception of the lagged outcome and the strata fixed effect. Second, we re-estimate the LATE as in (2) using the general variable for VODP2 activities take-up. Results are reported in Panel A and B in Tables C.2 to C.8 in Appendix C. Overall, results hold under the different alternative specifications.

## 4.2 Impact on oilseed adoption

We report the estimated impact on farmers' adoption of oilseed crops in Table 6. In each column, we report both the ITT estimate and the LATE in the third line. First, we find that the program has a positive and significant effect on the adoption of oilseed crops. In column 1, we find that the exposure to the program increases the probability of adopting by 3.7 percentage point. The point estimate is statistically significant at the 1% level, and corresponds to an increase of 15.10% relative the control group. This finding at the extensive margin is corroborated by the program impact at the intensive margin shown in column 2, where we show that share of land allocated to oilseed crops increases by 0.063 percentage points. Although this point estimate is relatively small, it is precisely estimated (significant at the 1% level) and translates to 17.63% increase relative to the control group. In column 3, we show the effect on the KLK index. The point estimate is 0.081 and significant at the 1% level whereas the p-value corrected for multiple hypothesis testing is 0.014. This confirms that the program unambiguously increased the adoption of oilseed crops at both the extensive and the intensive margins.

We report the LATE in the second row. In the last row, we show the F-statistic of the null hypothesis that the coefficient of the excluded instrument is zero in the first stage regression, which is always largely rejected (Stock and Yogo, 2005). We report the findings of the first stage of model 2 in Table C.1. As expected, treatment assignment is a strong predictor of project activities take-up.

The LATE point estimates in Table 6 indicate that among households who were induced to take-up VODP2 in wake of the treatment, adoption rates of oilseed crops increased by 47.7 percentage points, whereas the share of land allocated to oilseed crops increased by 8.03% percentage

---

<sup>24</sup>The z-scores are calculated by subtracting the control group mean and dividing by the control group standard deviation.

points. Both coefficients are tightly estimated, being statistically significant at the 1% level. These estimates are more than 10 times the estimates of the ITT, suggesting that adoption was definitively higher among the households that took-up the treatment than the average household in a treated sub-county.

### 4.3 Adoption by profitability expectations

In what follows, we formally explore the extent to which farmers' *ex-ante* beliefs as well as perceptions on profitability (i.e. price and yield expectations) influence adoption decisions. We start by examining the manner in which the response to the treatment varies with beliefs in prices and yields at baseline. Subsequently, we demonstrate how the information delivered by the program contributes to shape (and revise) farmers' expectations on oilseed price and yields.

We first test heterogeneity treatment effects by including expectations at baseline along with its interaction with treatment in equation (1) above.

Estimates are reported in Table 7. In columns 1-4, we investigate the extent to which the treatment effect on adoption varies with heterogeneity in expected profitability at baseline. We measure the expected profitability by expected yields in column 1 and expected price in column 3. Additionally, we consider the wedge (i.e. percent deviation) between expected yields and observed yields (column 2) and expected price and retail price (column 4).

The point estimate of the interaction term is negative and statistically significant at the 5% level in column 1, thus indicating that the treatment effect is larger the lower expected yield is at baseline. To picture, with the treatment effect on the vertical axis and expected yield on the horizontal axis, the treatment effect reaches its highest level when the expected yield is zero and the treatment effect is found to be downward slopping. Price expectations heterogeneity has similar effects, as seen in column 3. Again, the interaction term is negative and statistically significant at the 1% level. Hence, farmers who underestimated oilseed price at baseline were more likely to alter their adoption decisions due to the treatment. We report heterogeneous effects across yield and price wedge in columns (2) and (4). In column (2), the coefficient of the interaction term is statistically indistinguishable from zero: it is small in magnitude and noisily estimated. Conversely, the interaction term in column 4 is positive and statistically significant at the 1% level. Hence, the treatment was found to be more effective when it enabled farmers to correct their perception on prices when they were correcting their perceptions relating to fields

Overall, the estimates point to the role played by expectations, in particular by price expectations, in terms of adoption decisions. The value of extension service programs seems to be larger for those with low expectations/higher wedge (with low information), as compared to better informed farmers. Put differently, price uncertainty may significantly deter cash crop adoption.

A potential concern is that profitability expectations are picking up the effects of some unobservable variables. Specifically, it is plausible that treatment effects vary with a certain number of variables that are measured at baseline, whereas profitability expectations are correlated with them. For example, farmers who are more familiar with oilseeds or better-off may be better informed about oilseed yield and price. In order to address this pressing concern, we explore heterogeneity in farmers' familiarity with oilseed (i.e. having ever cultivated an oilseed crop and having grown it at the baseline) as well as household wealth (i.e. total land and wealth index) at baseline. These estimates have been reported in columns 5-8. The coefficient of the interactions is found to be consistently small in magnitude and not statistically significant at any reasonable level. While this is not a definitive proof, it is reassuring to notice that the treatment effects do not change with some of the variables that may be correlated with profitability expectations.

In Table C.9, in the Appendix, it is shown that the results are robust to the simultaneous inclusion of all the dimensions of heterogeneity considered in Table 7, to account for possible cross-correlations among the dimensions of heterogeneity. The results elucidated in Table 7 above provide some strong evidence that the treatment effect on adoption is indeed heterogeneous across price and yield expectations of farmers at baseline.

As discussed in subsection 3.3, data on expectations are notoriously difficult to elicit. A consequence is that expectation variables usually contain a fraction of missing values. We now briefly discuss how missing values might have affected our study's results. In Appendix A we analyze the determinants of missing values in expected yield and price at baseline. In addition, we demonstrate that the probability of missing values for the expectation measures is not correlated with the treatment; hence, missing values do not threaten the internal validity of the exercise. However, some systematic patterns do arise when we look at the correlates. As expected, farmers who did not adopt oilseeds at the baseline, were less educated and poorer were less likely to report yield and price expectations. As a robustness check, we repeat the analysis in Table 7 on the restricted sample of households reporting expectations at both baseline and endline<sup>25</sup>. The results are displayed in Table C.10. The estimates of yield expectations and wedge in columns 1 and 2 decrease in size and are less precise, but still have the same signs as shown in Table 7. The estimates for price expectations and wedge in columns 3 and 4 are broadly similar to those that are shown in Table C.9, both in terms of magnitude and size.

---

<sup>25</sup>This is the sample that has been used in the regressions in Tables 8 and 9.

#### 4.4 Change in profitability expectations

In the former section, we show that treatment effect on adoption is heterogeneous across farmers expectations and those with higher (price) wedge are more affected by the intervention. In order to understand whether expectations are a main mechanism for farmers' technology switch, in this section we test whether or not the extension service program has a causal impact in changing farmers expectations. More specifically, we report ITT and LATE results as above in Table 8, where the dependent variable is either yield or price expectations at the endline (columns 1 and 2). We do the same while using the yield and price wedge as outcome variables (columns 3 and 4). According to the results, the treatment has a causal impact in changing farmers' price expectations (columns 2 and 4) but does not have any effect on farmers' beliefs about crop yields. The ITT estimate in column 2 implies that farmers exposed to the treatment revised expected prices up by 145 UShs, statistically significant at the 10% level. The difference between the expected price and observed retail shrunk by 264 UShs. The ITT estimates and the LATE in columns 2 and 3 are found to be remarkably similar. This suggests that information on price is shared within the sub-county so that farmers in treated sub-counties ended up with the same information set whether or not they took-up the program.

We further explore farmers' revision of their expectations by explicitly considering farmers who over-estimated yields and prices from those who under-estimated at the baseline. Table 9 exhibits the estimation results. We focus mainly on price expectations in columns 1-3 since the estimates of yield expectations in columns 4-6 are imprecise. The dependent variable is the wedge in expected price, i.e. the percentage deviation from actual price, at endline (it is positive for respondents who under-estimate the market price and negative for those under-estimating it). In column 1, the treatment effect is estimated at -0.0441 and is found to be significant at the 1% level, which is equivalent to revising expected price up by 441 UShs. The coefficient of the interaction term (between wedge at baseline and treatment) is 0.0392 and significant at the 10%, thus indicating that the size of the revisions is sensitive to the magnitude of the initial error. In particular, the adjustment in response to the treatment, i.e. the reduction in endline price expectations wedge, is higher for farmers with a smaller baseline wedge in price expectations. Column 2 reveals that the treatment does not have any differential impact on the price wedge for farmers under vs over-estimating the market price at the baseline, on average. In column 3, we consider the triple interaction, by allowing the effect of the treatment to vary with the size of the wedge separately for those under and over-estimating the market price. We find that the entire treatment effect on the wedge reduction is attributed to farmers underestimating the market price. We interpret this result as evidence that farmers revise expectations in a logical way, which in turn, supports the

significant role of ex-ante farmers beliefs as a significant driver of farmers adoption choices.

As discussed earlier, due to non-response, the measures of oilseeds profitability expectations are subject to missing values. The above analysis is contingent on the sample of farmers with non-missing information in both waves. We relax the restriction by repeating the exercise excluding from the controls the outcome variable at the baseline. This, in turn, allows us to restrict the sample to non-missing observations only at the endline. Results, reported in Table C.11, are very similar to those in Table 8. Evidently, non-missing responses to the questions about oilseed profitability expectations are significantly correlated with the treatment assignment.<sup>26</sup> The simple fact of reporting could be viewed as an outcome of the program. This suggests that the interpretation of results in Tables 8 and 9 should be interpreted as conditional on reporting the information.

## 4.5 Impact on intermediate outcomes

In this section, we investigate the impact of the program on a set of intermediate outcomes, namely, the use of farm inputs (Table 10), labor allocation (Table 11) and access to markets and/or marketing strategies (Table 12). Starting with input use in Table 10, we reveal the effects of the program on fertilizers (columns 1-3), pesticides (4-6), improved seeds (7-8), and on an aggregate KLIK index. Our estimates indicate that the program increased expenses on fertilizers and the use of improved seeds. The former increased by 353 UShs, whereas the latter increased by 0.026 percentage points. The point estimates are significant at the 5% and the 1% levels, respectively. The estimated effects are large in relative terms, translating to 163% and 26.7% increase as compared to the control group. The effects on the remaining outcomes in this Table are imprecisely estimated, but are positive, with the exception of the effect on pesticide expenses, which is found to be negative, albeit very small in magnitude. However, the program's overall effect on these sets of outcomes is positive and significant, as reflected by the estimated impact on the KLIK index in column 9, although the effect is not robust to the multiple hypothesis correction.

In Table 11 we report the program effect on labor allocation. More specifically, we estimate the program effect on farm labor supply on oilseed crops. We consider total labor in column 1, labor from the household head, his spouse and other family members in columns 2-4, hired non-family labor in column 5, as well as a KLIK aggregate index in column 6. The program effects in column 1-4 are consistently positive, but imprecisely estimated. There seems to be an increased use of outside labor; the point estimate is 0.355 and statistically significant at the 1% level. The effect on the KLIK index is also found to be positive and significant at the 10% level, but does not survive to the multiple hypothesis correction. When taking these estimates at face value, it appears that

---

<sup>26</sup>About 50% and 44% of treated and control farmers respond to the questions on yield expectations at the endline, respectively. The difference is significant at 1% confidence level. Similar figures are applicable to price expectations.

the program prompted farmers to hire more outside labor. It can be speculated that the need to hire outside labor stems from the adoption of oilseed crops. The size of the LATE (4.447) provides some supportive evidence to that end. There are two possible ways in which adopting oilseed crop can lead to an increased demand for outside labor. First, it is plausible that oilseed crops are more labor intensive than the existing crops, so that reallocating land away from existing crops towards oilseeds leads to an increase in labor demand. Second, when it comes to cash crops, certain farming tasks may need to be performed at specific periods of the crop cycle in order to ensure quality. Family labor may then be insufficient in certain periods of the farming season, thus leading the household to hire outside labor.

The program's impacts on access to markets and marketing strategies are shown in Table 12. We find a positive and significant treatment effect on the likelihood of selling (any) oilseeds in column 1 and on bulk/group sales in column 3. The ITT point estimates imply an increase of 2.5 percentage point for oilseeds sales as well as an increase of 0.06 for bulk sales. Although these point estimates are small in magnitude, they translate into substantial relative changes. Compared to the control, sales of oilseeds increased by 19.7% and bulk sales increased by 73.3%. We do not detect any significant effect on buying seeds through the market (column 2), contacts with value chain actors (column 3) and sale to market actors. Although the effect on the aggregate KLIK index is positive, significant and strong enough to pass the multiple hypothesis correction, the figures indicate that the program was not successful at linking farmers to the market, particularly with the value chain actors.

#### **4.6 Impact on productivity and farm revenues**

Finally, we estimate the effects on farm productivity, total output and household income. We report the impact of the program on farm productivity in Table 13. We measure productivity by physical yields (kilograms per acre) in column 1 and the monetary value of harvest per acre in column 4 by applying market prices to the quantities harvested. The point estimates are all found to be positive and statistically significant at the 5% level. The figures imply an increase of 26.89% in physical yields and 30.03% in yield value in comparison to the control group. Those are substantial gains, given the relatively short duration of the program and its variability in intensity. While the productivity gains are noticeable, we do not detect any significant effect on total harvest in column 2, harvest value in column 3, revenues from sales in column 5 as well as on the aggregate KLIK index in column 6. Although the estimates coefficients for these outcomes variables are all positive and in fact large in some cases, they are imprecisely estimated, thus making it difficult to rule out the null hypothesis that the program did not have any effect.

In Table 14, we investigate the impact of the program on a number of household level out-

comes, including total cultivated land, total household farm labor supply, household farm income, household income from labor sale, and wealth index. In addition, we report the estimate on the aggregate KKK index based on all these variables. Overall, the effects are imprecisely estimated, with the exception of sesame in column 5. We do not detect any significant effect on any of these household level outcome variables.

The results presented suggest that VODP2 has positive and statistically significant effects on cash crops adoption, use of inputs and productivity; however, the magnitudes are modest. On the other hand, the program has little effect on final outcomes in terms of revenues and well-being. These findings are in consonance with the results of several studies which did not find a significant rise in household income following an increase either in technology adoption (e.g. Beaman et al., 2013) or in the short run (e.g. Carter et al., 2014).

## 5 Conclusion

In this paper, we use a randomized controlled experiment in order to estimate the causal impact of a large-scale agricultural extension program run by the Government of Uganda on oilseed adoption and downstream outcomes among smallholder farmers in the country. The extension program is a bundled intervention that provides technical training and market information to farmers. The services are supplied by private providers through farmers groups and value chain nurturing. A critical component was supposed to provide farmers with technical information in addition to the correct information on the returns to oilseeds, including information on prices, yields and access to market.

We find that the program significantly increased the adoption of oilseed crops. At endline, adoption rates and share of land allocated to oilseed crops increased by 15.10% and 17.63% in the treated group when compared with the control group. These gains are indeed substantial, given that the program lasted less than a calendar year. We further underpin the importance of price expectations on farmers' adoption decisions. In particular, we find that farmers with low expected price of oilseed crops at baseline are more likely to adopt. Specifically, farmers whose expected price at baseline is smaller than the actual market price are more likely to adopt after being treated. Moreover, by finding that the program updates and corrects farmers' beliefs about profitability, particularly by reducing the wedge in expected price, we show that price uncertainty rather than technical mis-information plays a key role in farmers' adoption choices.

Taken together, our findings confirm that market information is key in farming choices, even though higher uptake is not coupled with better performance in terms of short-term household well-being. This is particularly supported by the modest impact we find on farm household revenues



and income. If lack of information and ambiguity about future crop profitability – especially in terms of expected price – may restrict adoption rates, lack of proper farmers’ integration into the downstream part of the value-chain may pose a threat to the ‘follow-through stages’ in oilseed development. Put succinctly, disadoption may occur if expected returns to technology adoption are not realized. Impact evaluations are usually of specific interventions in a specific context. It is possible that these results cannot be generalized to different contexts. However, they contribute to the provision of systematic and causal evidence on the efficacy of national and international efforts to spur agricultural technology innovation and cash crops in Uganda, where agriculture continues to be one of the mainstays of its economy, engaging two thirds of the active population.

## References

- Aker, Jenny C.**, “Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger,” *American Economic Journal: Applied Economics*, July 2010, 2 (3), 46–59.
- Ambler, Kate, Alan de Brauw, and Susan Godlonton**, “Agriculture support services in Malawi: Direct effects, complementarities, and time dynamics,” *IFPRI Discussion Paper 01725*, 2018.
- Anderson, Jock R. and Gershon Feder**, “Agricultural Extension: Good Intentions and Hard Realities,” SSRN Scholarly Paper ID 873726, Social Science Research Network, Rochester, NY February 2008.
- Angrist, Joshua D, Guido W Imbens, and Donald B Rubin**, “Identification of causal effects using instrumental variables,” *Journal of the American statistical Association*, 1996, 91 (434), 444–455.
- Arouna, Aminou, Jeffrey D Michler, and Jourdain C Lokossou**, “Contract Farming and Rural Transformation: Evidence from a Field Experiment in Benin,” Working Paper 25665, National Bureau of Economic Research March 2019.
- Ashour, Maha, Daniel Orth Gilligan, Jessica Blumer Hoel, and Naureen Iqbal Karachiwalla**, “Do Beliefs About Herbicide Quality Correspond with Actual Quality in Local Markets? Evidence from Uganda,” *The Journal of Development Studies*, 2019, 55 (6), 1285–1306.
- Ashraf, Nava, Xavier Giné, and Dean Karlan**, “Finding missing markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in Kenya,” *American Journal of Agricultural Economics*, 2009, 91 (4), 973–990.
- Attanasio, Orazio and Vincenzo di Maro**, “Income Expectations in Rural Mexico: Measurement and Models,” 2008.
- Attanasio, Orazio P.**, “Expectations and Perceptions in Developing Countries: Their Measurement and Their Use,” *American Economic Review*, May 2009, 99 (2), 87–92.
- **and Katja M. Kaufmann**, “Education choices and returns to schooling: Mothers’ and youths’ subjective expectations and their role by gender,” *Journal of Development Economics*, July 2014, 109, 203–216.

- Barrett, Christopher and Paul Dorosh**, “Farmers’ Welfare and Changing Food Prices: Non-parametric Evidence from Rice in Madagascar,” *American Journal of Agricultural Economics*, 1996, 78 (3), 656–669.
- Barrett, Christopher B, Asad Islam, Abdul Malek, Debayan Pakrashi, and Ummul Ruthbah**, “The Effects of Exposure Intensity On Technology Adoption and Gains: Experimental Evidence from Bangladesh on the System of Rice Intensification,” 2018.
- Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry**, “Profitability of Fertilizer: Experimental Evidence from Female Rice Farmers in Mali,” *The American Economic Review*, 2013, 103 (3), 381–386.
- , – , – , and – , “Self-selection into credit markets: Evidence from agriculture in Mali,” Technical Report, National Bureau of Economic Research 2014.
- Bellemare, Marc F. and Jeffrey R. Bloem**, “Does contract farming improve welfare? A review,” *World Development*, 2018, 112, 259 – 271.
- Bold, Tessa, Kayuki C. Kaizzi, Jakob Svensson, and David Yanagizawa-Drott**, “Lemon Technologies and Adoption: Measurement, Theory and Evidence from Agricultural Markets in Uganda\*,” *The Quarterly Journal of Economics*, 03 2017, 132 (3), 1055–1100.
- Bonan, Jacopo and Laura Pagani**, “Junior Farmer Field Schools, Agricultural Knowledge and Spillover Effects: Quasi-Experimental Evidence from Northern Uganda,” *The Journal of Development Studies*, November 2018, 54 (11), 2007–2022.
- Burke, Marshall, Lauren Falcao Bergquist, and Edward Miguel**, “Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets\*,” *The Quarterly Journal of Economics*, 12 2018, 134 (2), 785–842.
- Carter, Michael R, Rachid Laajaj, and Dean Yang**, “Subsidies and the persistence of technology adoption: Field experimental evidence from Mozambique,” Technical Report, National Bureau of Economic Research 2014.
- Cole, Shawn A. and A. Nilesh Fernando**, “‘Mobile’izing Agricultural Advice: Technology Adoption, Diffusion and Sustainability,” Technical Report 13-047, Harvard Business School March 2018.
- de Janvry, A., E. Sadoulet, and T. Suri**, “Chapter 5 - Field Experiments in Developing Country Agriculture,” in Abhijit Vinayak Banerjee and Esther Duflo, eds., *Handbook of Economic*

*Field Experiments*, Vol. 2 of *Handbook of Economic Field Experiments*, North-Holland, January 2017, pp. 427–466.

**Delavande, Adeline and Basit Zafar**, “University Choice: The Role of Expected Earnings, Nonpecuniary Outcomes, and Financial Constraints,” *Journal of Political Economy*, December 2018, pp. 000–000.

– **and Hans-Peter Kohler**, “HIV/AIDS-related Expectations and Risky Sexual Behaviour in Malawi,” *The Review of Economic Studies*, January 2016, 83 (1), 118–164.

– **, Xavier Giné, and David McKenzie**, “Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design?,” *Journal of Applied Econometrics*, 2011, 26 (3), 479–497.

**Deutschmann, Joshua W., Maya Duru, Kim Siegal, and Emilia Tjernström**, “Can Smallholder Extension Transform African Agriculture?,” NBER Working Papers 26054, National Bureau of Economic Research, Inc 2019.

**Duflo, Esther, Michael Kremer, and Jonathan Robinson**, “Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya,” *American Economic Review*, 2011, 101 (6), 2350–90.

**Fafchamps, Marcel**, “Cash Crop Production, Food Price Volatility, and Rural Market Integration in the Third World,” *American Journal of Agricultural Economics*, 1992, 74 (1), 90–99.

– **and Bart Minten**, “Impact of SMS-Based Agricultural Information on Indian Farmers,” *The World Bank Economic Review*, January 2012, 26 (3), 383–414.

**Feder, Gershon, Lawrence J. Lau, and Roger H. Slade**, “Does Agricultural Extension Pay? The Training and Visit System in Northwest India,” *American Journal of Agricultural Economics*, August 1987, 69 (3), 677–686.

– **, Richard E Just, and David Zilberman**, “Adoption of agricultural innovations in developing countries: A survey,” *Economic development and cultural change*, 1985, 33 (2), 255–298.

**Filmer, Deon and Lant H Pritchett**, “Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India,” *Demography*, 2001, 38 (1), 115–132.

- Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein**, “Learning through noticing: Theory and evidence from a field experiment,” *The Quarterly Journal of Economics*, 2014, 129 (3), 1311–1353.
- Hochberg, Yosef**, “A sharper Bonferroni procedure for multiple tests of significance,” *Biometrika*, 1988, 75 (4), 800–802.
- Imbens, Guido W**, “Instrumental Variables: An Econometrician’s Perspective,” *Statistical Science*, 2014, 29 (3), 323–358.
- and **Joshua D Angrist**, “Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, 1994, 62 (2), 467–475.
- Jack, B Kelsey**, “Private information and the allocation of land use subsidies in Malawi,” *American Economic Journal: Applied Economics*, 2013, 5 (3), 113–35.
- Jensen, Robert**, “The Digital Divide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector,” *The Quarterly Journal of Economics*, January 2007, 122 (3), 879–924.
- , “The (Perceived) Returns to Education and the Demand for Schooling,” *The Quarterly Journal of Economics*, May 2010, 125 (2), 515–548.
- Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry**, “Agricultural Decisions after Relaxing Credit and Risk Constraints,” *The Quarterly Journal of Economics*, May 2014, 129 (2), 597–652.
- Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz**, “Experimental analysis of neighborhood effects,” *Econometrica*, 2007, 75 (1), 83–119.
- Kondylis, Florence, Valerie Mueller, and Jessica Zhu**, “Seeing is believing? Evidence from an extension network experiment,” *Journal of Development Economics*, 2017, 125, 1–20.
- Maertens, Annemie, Hope Michelson, and Vesall Nourani**, “How Do Farmers Learn From Extension Services?: Evidence From Malawi,” 2018.
- Maffioli, Alessandro, Diego Ubfal, Gonzalo Vazquez-Bare, and Pedro Cerdan-Infantes**, “Improving technology adoption in agriculture through extension services: evidence from Uruguay,” *Journal of Development Effectiveness*, 2013, 5 (1), 64–81.

- Magruder, Jeremy R.**, “An Assessment of Experimental Evidence on Agricultural Technology Adoption in Developing Countries,” *Annual Review of Resource Economics*, 2018, 10 (1), 299–316.
- Manski, Charles F.**, “Measuring Expectations,” *Econometrica*, 2004, 72 (5), 1329–1376.
- McKenzie, David, John Gibson, and Steven Stillman**, “A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad?,” *Journal of Development Economics*, May 2013, 102, 116–127.
- Meiburg, Charles O and Karl Brandt**, “Agricultural Productivity in the United States: 1870-1960,” *Food Research institute studies*, 1962, 3 (1387-2016-116176), 63.
- Minten, Bart and Steven Kyle**, “The effect of distance and road quality on food collection, marketing margins, and traders’ wages: evidence from the former Zaire,” *Journal of Development Economics*, 1999, 60 (2), 467 – 495.
- Oliva, Paulina, B. Kelsey Jack, Samuel Bell, Elizabeth Mettetal, and Christopher Severen**, “Technology Adoption Under Uncertainty: Take-Up and Subsequent Investment in Zambia,” *The Review of Economics and Statistics*, March 2019, pp. 1–42.
- Stifel, David and Bart Minten**, “Isolation and agricultural productivity,” *Agricultural Economics*, 2008, 39 (1), 1–15.
- Stock, James H and Motohiro Yogo**, “Testing for weak instruments in Linear Iv regression,” in “Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg,” Cambridge University Press, 2005, pp. 80–108.
- Sunding, David and David Zilberman**, “The agricultural innovation process: research and technology adoption in a changing agricultural sector,” in Bruce L. Gardner and Gordon C. Rausser, eds., *Handbook of agricultural economics. Volume 1*, Handbooks in Economics, vol. 18, Amsterdam; London and New York: Elsevier Science, North-Holland., 2001, chapter 4, pp. 207–261.
- Suri, Tavneet**, “Selection and comparative advantage in technology adoption,” *Econometrica*, 2011, 79 (1), 159–209.
- Waddington, Hugh, Birte Snilstveit, Jorge Garcia Hombrados, Martina Vojtkova, Jock Anderson, Howard White, Hugh Waddington, Birte Snilstveit, Jorge Garcia Hombrados, Martina Vojtkova, Jock Anderson, and Howard White**, “Farmer Field Schools

for Improving Farming Practices and Farmer Outcomes in Low- and Middle-income Countries: A Systematic Review,” November 2012.

**Wiswall, Matthew and Basit Zafar**, “How Do College Students Respond to Public Information about Earnings?,” *Journal of Human Capital*, June 2015, 9 (2), 117–169.

**World Bank**, “World development report 2008: Agriculture for development,” *The International Bank for Reconstruction and Development, Washington, DC*, 2007.

**Yishay, Ariel Ben, Maria Jones, Florence Kondylis, and Ahmed Mushfiq Mobarak**, “Gender gaps in technology diffusion,” *Journal of Development Economics*, September 2019, p. 102380.

Figures and Tables

Figure 1: Map of the study areas

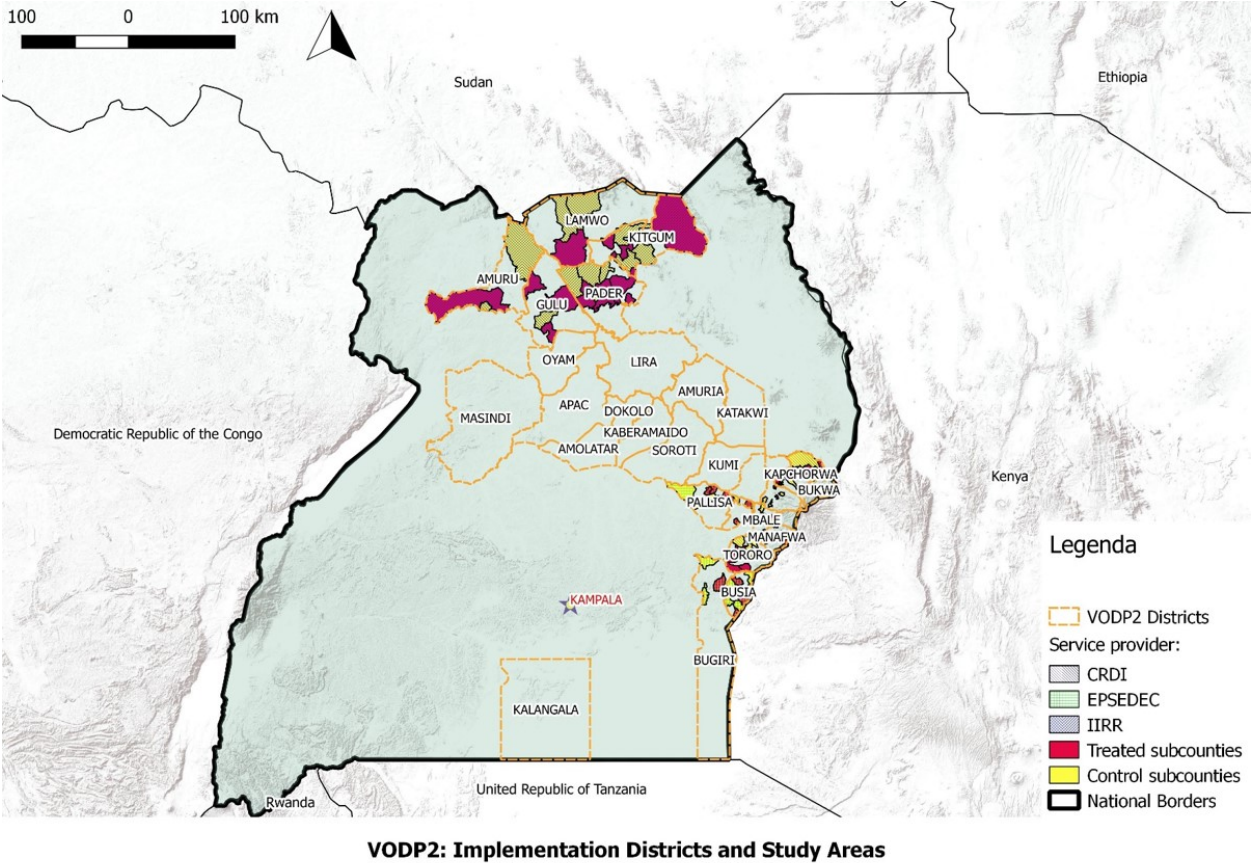


Figure 2: Study timeline

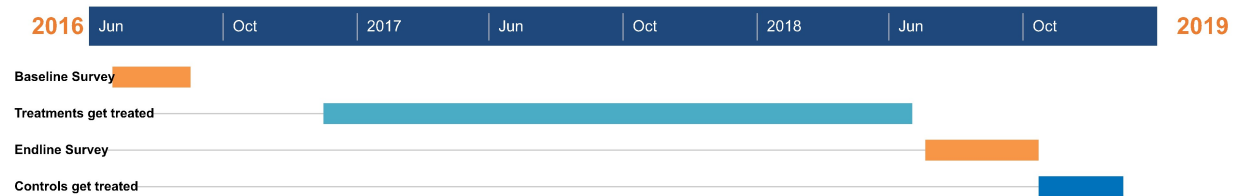




Table 1: Summary statistics and balance at household level

	(1)	(2)
	Control mean	ITT
<i>PANEL A: Respondent characteristics</i>		
HH head	0.604 (0.489)	0.00452 (0.0336)
Male	0.623 (0.485)	0.000 (0.0338)
Can read	0.748 (0.491)	-0.0260 (0.0253)
Can write	0.741 (0.495)	-0.0184 (0.0255)
No education	0.0960 (0.295)	0.0166 (0.0193)
Primary education	0.484 (0.500)	-0.0157 (0.0255)
Secondary education	0.379 (0.485)	-0.0111 (0.0269)
Above secondary education	0.0410 (0.198)	0.010 (0.009)
<i>PANEL B: HH level general outcomes</i>		
N. of plots cultivated	2.293 (1.245)	-0.107 (0.0777)
Total land	6.648 (10.07)	0.349 (0.923)
HH days of farm work	233 (132)	-6.356 (8.581)
Revenues from crop sale	133.6 (323)	-19.06 (22.46)
HH monthly labour income	23.56 (51.74)	1.640 (3.380)
Wealth index	-0.0159 (1.945)	0.0318 (0.102)

*Note:* This table reports household level summary statistics and assesses balance across groups using data from the baseline survey. Variables are described in Appendix B. Total cultivated land is expressed in acres, revenues from crop sale and HH labour income are in 10,000 Ush. Column 1 shows the summary statistics (mean and standard deviation) for the control group at baseline. Column 2 reports an Intention to Treat (ITT) estimate of the difference in means (and robust standard error clustered at the sub-county level) between the treatment and control group. The sample size is 2733. Asterisks denote statistical significance: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table 2: Oilseed-specific summary statistics and balance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Soyabean		Sunflower		Sesame		Groundnuts	
	Control mean	ITT	Control mean	ITT	Control mean	ITT	Control mean	ITT
<i>PANEL A: Adoption</i>								
Adoption	0.222 (0.416)	0.0236 (0.0429)	0.108 (0.310)	-0.00462 (0.0380)	0.210 (0.408)	0.0171 (0.0627)	0.437 (0.496)	0.00416 (0.0525)
Share of land	0.0309 (0.0784)	0.00470 (0.00746)	0.0256 (0.0935)	-0.00271 (0.0124)	0.0400 (0.102)	-0.00256 (0.0134)	0.0679 (0.109)	0.000173 (0.00995)
<i>PANEL B: Inputs</i>								
Fertilizer use	0.0344 (0.182)	-0.00592 (0.0101)	0.00659 (0.0810)	-0.00148 (0.00303)	0.0110 (0.104)	-0.00149 (0.00449)	0.0491 (0.216)	-0.0133 (0.0105)
Fertilizer quantity	367.1 (4826)	-171.7 (165.0)	60.65 (1739)	-11.14 (60.26)	6.214 (205.3)	-6.212 (5.537)	359.6 (4107)	300.5 (415.5)
Fertilizer expense	0.0382 (0.511)	-0.00751 (0.0214)	0.0228 (0.473)	-0.0127 (0.0141)	0.000621 (0.0205)	-0.000621 (0.000554)	0.0608 (0.828)	0.0133 (0.0471)
Pesticide use	0.0278 (0.165)	-0.00883 (0.00858)	0.00879 (0.0934)	-0.00148 (0.00456)	0.00366 (0.0604)	0.000723 (0.00228)	0.0542 (0.227)	-0.0220* (0.0130)
Pesticide quantity	369.7 (4821)	-159.2 (166.1)	60.73 (1740)	-11.05 (60.44)	6.171 (204.7)	-6.170 (5.502)	359.9 (4120)	260.0 (416.6)
Pesticide expense	0.0272 (0.386)	0.00171 (0.0253)	0.0100 (0.210)	-0.0100* (0.00555)	0.000382 (0.0136)	-0.000308 (0.000372)	0.0522 (0.499)	-0.0449*** (0.0140)
Improved seeds use	0.0872 (0.282)	0.0137 (0.0204)	0.0520 (0.222)	0.00427 (0.0196)	0.0520 (0.222)	0.00354 (0.0160)	0.182 (0.386)	0.00179 (0.0253)
Seeds expense	0.0547 (0.345)	0.0156 (0.0215)	0.0485 (0.355)	0.000649 (0.0265)	0.0198 (0.205)	0.00180 (0.0106)	0.204 (1.047)	0.0641 (0.0642)
<i>PANEL C: Labour supply (Days of work)</i>								
By all	5.493 (13.51)	1.349 (1.374)	3.492 (13.13)	-0.504 (1.426)	6.937 (17.33)	-0.598 (2.152)	12.44 (19.58)	0.267 (1.845)
By head	3.116 (8.281)	0.860 (0.855)	1.895 (7.866)	-0.292 (0.781)	4.191 (11.48)	-0.771 (1.273)	7.907 (13.38)	0.178 (1.330)
By spouse	3.205 (9.124)	0.979 (0.929)	1.902 (8.230)	-0.331 (0.727)	3.952 (11.44)	-0.698 (1.200)	8.123 (13.89)	-0.573 (1.161)
By other in the HH	1.680 (6.307)	0.353 (0.460)	0.668 (3.375)	-0.0580 (0.313)	1.503 (5.967)	-0.172 (0.465)	4.156 (10.58)	0.430 (0.782)
By other outside HH	0.543 (3.317)	-0.0563 (0.156)	0.341 (2.213)	-0.0345 (0.195)	0.422 (2.830)	-0.0870 (0.134)	1.617 (6.404)	-0.0288 (0.325)

*Note:* This table reports household level summary statistics for oilseeds and assesses balance across groups using data from the baseline survey. Odd columns show oilseed-specific summary statistics (mean and standard deviation) for the control group at baseline for soyabean, sunflower, sesame and groundnuts, respectively. Even columns report an Intention to Treat (ITT) estimate of the difference in means (and robust standard error clustered at the sub-county level) between the treatment and control group. Variables are described in Appendix B. Quantities of fertilizer and pesticide are expressed in Kg, expenses for fertilizer and pesticide are in 10,000 Ush. The sample size is 2733. Sample size for expectations varies and is described in detail in appendix A. Asterisks denote statistical significance: \*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01.

Table 2: Oilseed-specific summary statistics and balance (cont.)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Soyabean		Sunflower		Sesame		Groundnuts	
	Control mean	ITT	Control mean	ITT	Control mean	ITT	Control mean	ITT
<i>PANEL D: Market access</i>								
Any sale	0.143 (0.350)	0.0348 (0.0336)	0.0842 (0.278)	-0.00676 (0.0326)	0.123 (0.329)	0.0122 (0.0407)	0.281 (0.450)	-0.0189 (0.0414)
Seeds bought on the mkt	0.106 (0.308)	-0.00462 (0.0180)	0.0513 (0.221)	0.00647 (0.0142)	0.0447 (0.207)	-0.00668 (0.0105)	0.114 (0.318)	-0.00171 (0.0174)
Harvest sold in bulk/group	0.00586 (0.0764)	0.0102* (0.00537)	0.00806 (0.0894)	0 (0.00645)	0.00513 (0.0715)	-0.00147 (0.00348)	0.00879 (0.0934)	0 (0.00425)
Contacts with value chain actors	0.0916 (0.289)	-0.00970 (0.0241)	0.0821 (0.275)	-0.00310 (0.0204)	0.0432 (0.203)	0.0123 (0.0116)	0.0315 (0.175)	0.00432 (0.0103)
Sale to mkt actors	0.0300 (0.171)	0.0109 (0.0103)	0.0315 (0.175)	-0.00226 (0.0160)	0.0293 (0.169)	-0.00737 (0.0112)	0.0579 (0.234)	0.00207 (0.0124)
<i>PANEL E: Production and productivity</i>								
Productivity	0.592 (6.401)	0.262 (0.484)	0.323 (1.679)	0.387 (0.369)	0.367 (4.024)	0.139 (0.304)	1.706 (11.58)	-0.413 (0.481)
Harvest	0.388 (5.967)	0.371 (0.464)	0.625 (4.230)	0.189 (0.484)	0.500 (7.001)	-0.036 (0.335)	1.408 (20.19)	-0.334 (0.589)
Harvest value	8.673 (129.2)	9.160 (10.89)	7.884 (54.14)	1.659 (5.770)	21.68 (302.4)	-1.565 (14.48)	111.9 (944.3)	-4.194 (54.61)
Harvest value/acre	14.17 (179.2)	9.748 (14.35)	4.145 (25.56)	4.644 (5.050)	15.22 (168.2)	11.76 (15.70)	213 (2823)	-97.32 (134.6)
Sale revenues	2.204 (10.79)	0.677 (0.766)	4.775 (47.41)	-1.847 (3.132)	2.763 (11.75)	-0.598 (1.149)	9.070 (81.27)	-2.939 (2.888)
<i>PANEL F: Profitability expectations</i>								
Expected yield	5.312 (3.751)	0.389 (0.326)	7.382 (9.543)	0.295 (1.260)	3.939 (2.886)	-0.0271 (0.289)	7.414 (7.004)	0.101 (0.493)
Expected price	0.207 (0.0915)	-0.00452 (0.00858)	0.119 (0.0735)	0.00538 (0.00892)	0.859 (13.03)	-0.548 (0.529)	0.389 (0.502)	-0.0184 (0.0282)
Yield wedge	-5.280 (6.167)	0.00617 (0.871)	-2.018 (4.360)	0.00355 (0.582)	-2.772 (3.296)	-0.102 (0.471)	-2.707 (3.690)	-0.0785 (0.310)
Price wedge	0.399 (0.265)	0.0132 (0.0249)	0.316 (0.421)	-0.0308 (0.0511)	-0.314 (19.93)	0.839 (0.810)	0.415 (0.837)	0.0146 (0.0641)

*Note:* This table reports household level summary statistics for oilseeds and assesses balance across groups using data from the baseline survey. Odd columns show oilseed-specific summary statistics (mean and standard deviation) for the control group at baseline for soyabean, sunflower, sesame and groundnuts, respectively. Even columns report an Intention to Treat (ITT) estimate of the difference in means (and robust standard error clustered at the sub-county level) between the treatment and control group. Variables are described in Appendix B. Productivity is in quintal/acre, harvest is in quintal, harvest value is in 10,000 Ush, sale revenues is in 10,000 Ush. Expected yield are expressed in quintal/acre, price expectations are 10,000 Ush per Kg, yield and price wedge are percentage deviations from actual values. The sample size is 2733. Sample sizes for yield (price) expectations are 1588, 792, 1095, 2224 (1619, 744, 1128, 2257) for soyabean, sunflower, sesame and groundnuts, respectively. Asterisks denote statistical significance: \*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01.

Table 3: Determinants of oilseed profitability expectations at baseline

	(1)	(2)	(3)	(4)	(5)	(6)
	Expected price			Expected yield		
HH head	0.0308 (0.0355)	0.0332 (0.0351)	0.0344 (0.0345)	1.690*** (0.523)	1.673*** (0.520)	1.901*** (0.397)
Male	-0.137 (0.108)	-0.140 (0.111)	-0.135 (0.103)	-1.632*** (0.568)	-1.604*** (0.568)	-1.931*** (0.446)
Can read	0.0315 (0.0423)	0.0280 (0.0418)	0.0471 (0.0511)	-0.308 (0.496)	-0.368 (0.468)	-0.427 (0.488)
Can write	-0.0518 (0.0442)	-0.0466 (0.0415)	-0.0528 (0.0504)	0.590 (0.525)	0.620 (0.496)	0.647 (0.517)
No education	0.431 (0.457)	0.425 (0.449)	0.452 (0.472)	-1.591* (0.892)	-1.701* (0.877)	-1.838** (0.883)
Primary education	-0.0652 (0.0732)	-0.0718 (0.0810)	-0.0615 (0.0724)	-0.818 (0.850)	-0.914 (0.835)	-0.964 (0.844)
Secondary education	-0.0462 (0.0606)	-0.0523 (0.0680)	-0.0482 (0.0657)	-0.396 (0.846)	-0.466 (0.833)	-0.509 (0.837)
Wealth index	-0.0272 (0.0263)	-0.0283 (0.0278)	-0.0317 (0.0297)	0.150*** (0.0534)	0.120** (0.0538)	0.128** (0.0524)
Total land	0.000258 (0.000917)	9.54e-05 (0.000811)	-0.000235 (0.000613)	0.0130** (0.00570)	0.0106* (0.00582)	0.0117** (0.00571)
N. of plots cultivated	0.110 (0.118)	0.108 (0.115)	0.114 (0.121)	0.151 (0.115)	0.121 (0.113)	0.0480 (0.0942)
Oilseed baseline adoption		0.105 (0.126)			1.076*** (0.201)	
KLK for adoption			-0.0644 (0.0496)			0.376*** (0.126)
KLK for input use			-0.0150 (0.0176)			0.0312 (0.0855)
KLK for labour allocation			0.201 (0.186)			0.140 (0.146)
KLK for productivity			-0.0195 (0.0212)			0.491** (0.225)
KLK for mkt linkages			-0.121 (0.116)			0.405* (0.211)
Constant	-0.0225 (0.303)	-0.0614 (0.348)	-0.0270 (0.307)	5.958*** (1.101)	5.596*** (1.100)	6.265*** (1.098)
Observations	5,391	5,391	5,287	5,341	5,341	5,242
R-squared	0.005	0.006	0.007	0.104	0.117	0.132
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes

*Note:* The table shows regressions with price and yield expectations as dependent variables in column 1-3 and 4-6, respectively. The model is similar to the one shown in equation 1. Regressors are considered at the baseline. The model also includes dummies for the assigned service providers. Variables are described in Appendix B. Total land is in acres. The sample comprises non-attriter farmers reporting expectations at the baseline. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 4: Internal and external cross-validation of profitability expectations

	(1)	(2)	(3)	(4)	(5)
	Expectations about population's		Actual yield		Actual price
	Yield	Price	Survey	LSMS	
Expected yield	1.742* (1.030)		0.134*** (0.0395)	0.00939** (0.00369)	
Expected price		4.352*** (0.875)			0.000181 (0.000233)
Constant	436.0*** (31.20)	1,487*** (100.6)	2.103 (1.660)	2.068*** (0.0980)	1.379*** (0.0386)
Observations	3,585	3,651	5,342	5,341	5,392
R-squared	0.081	0.360	0.020	0.539	0.468
Controls	No	No	No	No	No
Strata FE	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes

*Note:* Columns 1 and 2 report correlations between own yield and price expectations and expectations about population's yield and price, respectively. All variables are taken at baseline. The sample is composed by non-attriter farmers reporting both own and population's yield and price expectations at the baseline, respectively. Columns 3 and 4 report correlation between own yield and price expectations and actual realizations. All variables are taken at baseline. Actual yields at the farmer level as measured in the survey are used in column 3, while actual yields from LSMS are used in column 4. Actual prices, in column 5, are taken from Trade-AGMIS. The sample is composed by non-attriter farmers reporting own yield and price expectations at the baseline. Strata are districts. The Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 5: Project take-up

	(1) Control mean	(2) ITT
VODP program take-up	0.241 (0.428)	0.475*** (0.0358)
Specific training/platform on soyabean	0.0298 (0.170)	0.178*** (0.0313)
Specific training/platform on sunflower	0.0149 (0.121)	0.0750*** (0.0221)
Specific training/platform on sesame	0.0266 (0.161)	0.0499** (0.0190)
Specific training/platform on groundnut	0.0188 (0.136)	0.0167 (0.0151)
N. of visits to learning platform	0.469 (2.539)	1.154*** (0.194)
Received advice on new seed varieties	0.380 (0.486)	0.216*** (0.0335)
Received advice on oilseed production	0.332 (0.471)	0.289*** (0.0378)
Received advice on farming techniques	0.599 (0.490)	0.167*** (0.0320)
Received advice on market activities	0.462 (0.499)	0.139*** (0.0328)
Received free inputs (any)	0.212 (0.409)	0.00371 (0.0323)
Received free seeds (any)	0.00235 (0.0485)	0.00317 (0.00258)

*Note:* This table reports household level summary statistics and assesses balance across groups using data from the endline survey. Variables are described in Appendix B. Column 1 shows the summary statistics (mean and standard deviation) for the control group at baseline. Columns 2 reports an Intention to Treat (ITT) estimate of the difference in means (and robust standard error clustered at the sub-county level) between the treatment and control group. The sample includes non-attriter farmers (N= 2544). Asterisks denote statistical significance: \*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01.

Table 6: Program impact on oilseed adoption

	(1)	(2)	(3)
	Oilseed adoption	Share of land cultivated with oilseed	KLK Index
<b>ITT</b>	0.0370*** (0.0112)	0.00626*** (0.00218)	0.0811*** (0.0252) [0.014]
<b>LATE</b>	0.477*** (0.149)	0.0803*** (0.0284)	1.043*** (0.332)
Observations	10,172	10,172	10,172
Controls	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes
Control mean	0.245	0.0355	0
F stat excl restr	88.78	89.29	89

*Note:* The table shows ITT and LATE as presented in model 1 and 2 in the text, respectively. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Hochberg-corrected p-value in square bracket. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 7: Heterogeneous effects on oilseed adoption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Expected profitability			Familyarity		Wealth		
	Expected yield	Yield wedge	Expected price	Price wedge	Ever cultivated	Cultivated at baseline	Total land	Wealth index
Treat	0.0650*** (0.0189)	0.0290* (0.0168)	0.0581*** (0.0202)	0.0105 (0.0166)	0.0335*** (0.0126)	0.0379*** (0.0123)	0.0398*** (0.0107)	0.0370*** (0.0112)
Het	0.00419*** (0.00109)	0.00152 (0.00232)	0.00182*** (0.000163)	-0.00115*** (0.000123)	0.118*** (0.0169)	0.262*** (0.0262)	0.000517 (0.000317)	0.00301 (0.00365)
Treat X Het	-0.00368** (0.00152)	-0.00388 (0.00301)	-0.0963*** (0.0334)	0.0483*** (0.0145)	0.00491 (0.0225)	-0.000354 (0.0310)	-0.000427 (0.000576)	0.000216 (0.00538)
Constant	0.363*** (0.0610)	0.390*** (0.0604)	0.368*** (0.0578)	0.374*** (0.0575)	0.246*** (0.0391)	0.292*** (0.0382)	0.291*** (0.0376)	0.292*** (0.0375)
Observations	5,341	5,341	5,391	5,391	10,145	10,172	10,172	10,172
R-squared	0.104	0.103	0.105	0.105	0.196	0.185	0.185	0.185
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Note:* The table shows the heterogeneous effects of the program on oilseed adoption, by baseline values of variables reported in the headings. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Expected yield is expressed in quintal/acre, expected price in 10,000 Ush, total land is in acre. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .



Table 8: Program impact on profitability expectations

	(1)	(2)	(3)	(4)	(5)
	Oilseed				KLK Index
	Expected yield	Expected price	Wedge in		
			Expected yield	Expected price	
<b>ITT</b>	-0.103 (0.817)	0.0145* (0.00755)	-0.361 (1.361)	-0.0265** (0.0123)	0.0304 (0.0216) [0.328]
<b>LATE</b>	-1.066 (8.424)	0.145* (0.0758)	-3.753 (13.93)	-0.264** (0.124)	0.315 (0.219)
Observations	3,600	3,626	3,600	3,626	3,890
Controls	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes
Control mean	5.320	0.303	-3.283	0.481	-0.000531
F stat excl restr	52.45	54.50	52.49	54.51	56.67

*Note:* The table shows ITT and LATE as presented in model 1 and 2 in the text, respectively. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Expected yield is expressed in quintal/acre, expected price in 10,000 Ush. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Hochberg-corrected p-value in square bracket. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 9: Revision of expectations

	(1)	(2)	(3)	(4)	(5)	(6)
	Wedge in					
	Expected price			Expected yield		
Treat	-0.0441*** (0.0154)	-0.117 (0.0799)	-0.107 (0.0837)	-5.119 (4.939)	-0.363 (1.442)	-5.597 (5.538)
Wedge	4.20e-05 (7.85e-05)		-0.000160* (8.17e-05)	1.951 (1.625)		2.009 (1.680)
Treat X Wedge	0.0392* (0.0224)		0.0204 (0.0144)	-1.389 (1.648)		-1.438 (1.706)
I[Wedge>0]		0.0694** (0.0281)	0.0320 (0.0313)		1.485 (1.130)	-3.726 (4.617)
Treat X I[Wedge>0]		0.0936 (0.0837)	0.135* (0.0810)		-1.061 (1.731)	2.753 (4.313)
I[Wedge>0] X Wedge			0.0761** (0.0373)			-5.655 (4.222)
Treat X I[Wedge>0] X Wedge			-0.124* (0.0711)			6.230 (5.324)
Constant	0.450*** (0.0406)	0.382*** (0.0478)	0.383*** (0.0477)	-0.247 (4.960)	-10.21** (3.934)	0.320 (5.388)
Observations	3,626	3,626	3,626	3,600	3,600	3,600
R-squared	0.056	0.061	0.063	0.025	0.012	0.026
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes

*Note:* The table shows the heterogeneous effects of the program on wedge in price (columns 1 to 3) and in yield expectations at the endline (columns 4 to 6), by values of the wedge and whether the wedge is positive (underestimation) at the baseline. The variable  $I[Wedge > 0]$  is a dummy which takes value one for underestimation of the true value at the baseline and zero otherwise. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Expected yield is expressed in quintal/acre, expected price in 10,000 Ush. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 10: Program impact on use of inputs for oilseeds

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Fertilizer		Pesticide		Improved seeds		KLK Index		
	Use	Quantity	Expenses	Use	Quantity	Expenses	Use	Expenses	
<b>ITT</b>	0.00600 (0.00362)	0.258 (0.185)	0.0353** (0.0135)	0.00732 (0.00942)	0.0736 (0.157)	-0.00123 (0.0120)	0.0255*** (0.00953)	0.0346 (0.0595)	0.0420** (0.0203) [0.205]
<b>LATE</b>	0.0769 (0.0490)	3.362 (2.479)	0.460** (0.192)	0.0939 (0.120)	0.985 (2.082)	-0.0164 (0.159)	0.328*** (0.127)	0.448 (0.764)	0.539** (0.271)
Observations	10,172	9,873	9,898	10,172	9,564	9,569	10,172	9,861	10,172
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.0204	0.427	0.0215	0.0880	0.306	0.0545	0.0956	0.413	0.0174
F stat excl restr	90.10	91.74	88.44	90.44	84.54	84.37	89.20	88.47	89.74

*Note:* The table shows ITT and LATE as presented in model 1 and 2 in the text, respectively. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Fertilizer and pesticide quantities are expressed in Kg, while expenses are in 10,000 Ush. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Hochberg-corrected p-value in square bracket. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 11: Program impact on labour allocation

	(1)	(2)	(3)	(4)	(5)	(6)
	Days of work on oilseeds by					KLK Index
	All	Head	Spouse	Other in the HH	Other outside the HH	
<b>ITT</b>	0.666 (0.462)	0.317 (0.298)	0.309 (0.269)	0.101 (0.223)	0.345*** (0.120)	0.0495* (0.0254) [0.244]
<b>LATE</b>	8.590 (6.018)	4.094 (3.868)	3.977 (3.480)	1.302 (2.854)	4.447*** (1.615)	0.638* (0.335)
Observations	9,828	9,828	9,828	9,828	9,828	9,828
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	6.210	3.593	3.515	2.344	0.617	0
F stat excl restr	87.94	87.79	87.64	87.92	87.77	87.98

*Note:* The table shows ITT and LATE as presented in model 1 and 2 in the text, respectively. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Hochberg-corrected p-value in square bracket. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 12: Program impact on oilseed market linkages

	(1)	(2)	(3)	(4)	(5)	(6)
	Any sale	Seeds bought on the mkt	Harvest sold in bulk/group	Contacts with value chain actors	Sale to mkt actors	KLK Index
<b>ITT</b>	0.0256** (0.0101)	0.00653 (0.00736)	0.00575** (0.00246)	0.00582 (0.00656)	0.000692 (0.00396)	0.0369*** (0.0140) [0.036]
<b>LATE</b>	0.329** (0.128)	0.0840 (0.0918)	0.0740** (0.0323)	0.0749 (0.0835)	0.00889 (0.0505)	0.475*** (0.176)
Observations	10,172	10,172	10,172	10,172	10,172	10,172
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.130	0.0521	0.00784	0.0462	0.0263	0.0104
F stat excl restr	88.59	89.77	90.07	89.74	89.33	89.68

*Note:* The table shows ITT and LATE as presented in model 1 and 2 in the text, respectively. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Hochberg-corrected p-value in square bracket. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 13: Program impact on oilseeds productivity

	(1)	(2)	(3)	(4)	(5)	(6)
	Productivity	Harvest	Harvest value	Harvest value/acre	Revenues from sale	KLK Index
<b>ITT</b>	0.0831** (0.0416)	0.038 (0.077)	2.042 (2.256)	4.850** (2.308)	0.616 (0.638)	0.0364 (0.0252) [0.249]
<b>LATE</b>	1.070** (0.543)	0.487 (0.980)	26.25 (28.91)	62.34** (31.04)	7.908 (8.119)	0.468 (0.326)
Observations	10,172	10,172	10,172	10,172	10,172	10,172
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.309	0.376	13.96	16.15	3.739	0
F stat excl restr	89.44	89.82	89.78	89.84	89.97	89.82

*Note:* The table shows ITT and LATE as presented in model 1 and 2 in the text, respectively. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Productivity is expressed in quintal/acre, harvest in quintal, harvest value, harvest value per acre and revenues from sale are expressed in 10,000 Ush. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Hochberg-corrected p-value in square bracket. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 14: Program impact on welfare

	(1)	(2)	(3)	(4)	(5)	(6)
	Total land	HH days of farm work	Rev from crop sale	HH labour income	Wealth Index	KLK Index
<b>ITT</b>	0.0754 (0.516)	-3.796 (7.408)	15.30 (11.26)	-5.716 (20.83)	0.0566 (0.0626)	0.0110 (0.0256) [0.668]
<b>LATE</b>	0.161 (1.088)	-8.070 (15.46)	32.50 (23.65)	-12.15 (43.73)	0.120 (0.132)	0.0234 (0.0539)
Observations	2,537	2,457	2,543	2,543	2,543	2,543
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	7.421	218.7	131.5	139.4	0.0976	0.0197
F stat excl restr	273.1	262	275.1	276.3	276	276

*Note:* The table shows ITT and LATE as presented in model 3 and 4 in the text, respectively. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Total land is expressed in acres, revenues from crop sale and HH labour income are in 10,000 Ush. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Hochberg-corrected p-value in square bracket. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

## Appendix

### A Attrition and missing values

In column 1 of Table A.1 we assess the determinants of attrition. First, we find that the treatment and control group are not affected differently by attrition. This ensures that attrition does not represent a threat to the internal validity of the study. Second, none of the observable characteristics seem to significantly predict attrition, which appears idiosyncratic. This suggests that attrition did not significantly alter the initial composition of the sample, in terms of main observable characteristics. Hence, the external validity of the study (conditional on the eligibility criteria of the program) is not threatened by attrition.

Part of the analysis involving expectations about oilseeds profitability relies on restricted samples of individuals who answered these questions at the baseline. We analyse systematic patterns in non-response to these questions in the remaining part of Table A.1. Columns 2 and 3 report the determinants of missing values in price and yield expectations at baseline, respectively. The dependent variables take the value of one when missing values are observed at the baseline. We find that missing values at the baseline are not significantly correlated to treatment status. We also find that missing values are more likely when farmers did not adopt oilseeds at baseline, are less educated (the omitted education dummy is above secondary education) and poorer. The results suggest that the exercise in Table 7 are not affected by non-response bias, as far as treatment is concerned. Additionally, our estimation results remain virtually the same whether we control or not the variables listed in the Table.



Table A.1: Analysis of attrition and missing values

	(1) Attriter	(2) Missing in Expected yield	(3) Missing in Expected price
Treat	0.0135 (0.0111)	0.00303 (0.0134)	0.00353 (0.0142)
Oilseed baseline adoption	0.00409 (0.0109)	-0.406*** (0.0208)	-0.390*** (0.0208)
HH head	-0.0375 (0.0511)	-0.000426 (0.0411)	-0.00763 (0.0421)
Male	0.0513 (0.0509)	-0.00622 (0.0395)	0.00583 (0.0406)
Can read	0.0213 (0.0348)	-0.0548 (0.0405)	-0.0646 (0.0410)
Can write	-0.0106 (0.0374)	0.0655 (0.0407)	0.0701* (0.0417)
No education	-0.0294 (0.0320)	0.109*** (0.0294)	0.119*** (0.0306)
Primary education	-0.0349 (0.0290)	0.0536** (0.0262)	0.0669** (0.0262)
Secondary education	-0.0406 (0.0270)	0.0459* (0.0249)	0.0599** (0.0255)
Wealth index	-0.00261 (0.00308)	-0.00902*** (0.00302)	-0.00511 (0.00313)
Total land	-0.000142 (0.000276)	-4.94e-05 (0.000402)	5.89e-05 (0.000281)
N. of plots cultivated	0.00128 (0.00436)	-0.00429 (0.00412)	-0.00447 (0.00439)
Constant	0.0399 (0.0350)	0.480*** (0.0380)	0.449*** (0.0370)
Observations	2,733	10,172	10,172
R-squared	0.033	0.390	0.382
Strata FE	Yes	Yes	Yes
Crop FE	No	Yes	Yes

*Note:* In column 1 the dependent variable is a dummy equal to one when the observation is missing at the endline and zero otherwise. The model is similar to the one shown in equation 3. In columns 2 and 3, the dependent variables are dummies equal to one when values in expected yield and price at the baseline are missing, respectively. The model is similar to the one shown in equation 1. Regressors are considered at the baseline. The model also includes dummies for the assigned service providers. Variables are described in Appendix B. Total land is in acres. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

## B Variables description

In this section, a detailed description on how variables presented in Tables 1 and 2 is provided. In general, some variables are measured only twice, i.e. at baseline and endline. Other variables are aggregated across the two baseline seasons and the two endline seasons.

### *Respondent and household characteristics*

Respondent characteristics are dummy variables which take the value of one when the condition in the label is satisfied and zero otherwise. Household land is divided into plots, which can be further divided into parcels and cultivated with different crops. "N. of plots cultivated" indicates the amount of plots considered for agricultural purposes, while total land is the sum of their areas, regardless of the property status.<sup>27</sup> Both variables calculated as the average across seasons. "HH days of farm work" is the weighted average of days spent in farm work across household members in the two baseline seasons. "HH labour income" is the sum of farm and off-farm incomes earned within the household in a month, on average. "Wealth index" is the first principal component of a batteries of question on the ownership of assets and livestock. In particular, the following variables are included: livestock ownership (sheep, goats, pack animals, ducks/turkeys, chicken, donkeys and pigs) and other assets and devices (i.e. fridge, sewing machine, tv, phone, radio, water pot, wooden or metal bed, mattress, bicycle, motorcycle and batteries).

### *Oilseed production*

Adoption is a dummy variable which takes value one if farmer cultivated the specific crop in at least one of the reference seasons and zero otherwise. Share of land is the weighted average of land dedicated to the specific crop across the two seasons. Likewise, fertilizer, pesticide and improved seeds use are dummies taking value one if it occurred in at least one of the reference seasons and zero otherwise. Quantities and expenses are averages across reference seasons.

Labour supply by different household or non-household members is expressed in days of work dedicated to any phase of cultivation of the specific oilseed.

Variables for market access are all dummies. "Any sale" is equal to one if some quantity of the specific oilseed was sold in at least one of the reference seasons, and zero otherwise. "Seeds bought on the market" is equal to one if oilseeds are supplied from market sources (e.g. private trader in local or district markets) in at least one of the reference seasons, and zero otherwise. "Contacts with value chain actors" takes value one if farmers got in contact with seed buyers, processor, miller or trader for each specific oilseed, and zero otherwise. "Sale to market actors" is equal to one if the harvest is sold to market actors (non-local markets traders or millers)<sup>28</sup>, and zero otherwise.

---

<sup>27</sup>Plots dedicated to the following categories are not computed in the sum: bush, forest, natural pastures.

<sup>28</sup>If we also include traders at local markets results do not change.

"Harvest" is the average harvest (in quintals) of the specific crop across the two reference seasons. When the crop is not cultivated, harvest is set to zero. "Productivity" is the average of season-and-oilseed-specific productivity, calculated as harvest divided by the amount of land dedicated to that crop. When the crop is not cultivated, productivity is set to zero. Productivity is expressed in quintals/acre. To construct "Harvest value" we multiply the harvest times the unit price from the administrative source. Values are expressed in 10,000 Ush. We then calculate "harvest value/acre", as an alternative productivity measure, by dividing harvest by the average land dedicated to that crop (the numerator of "share of land"). Values are expressed in 10,000 Ush. "Sale revenues" is the average amount made out of oilseed sales in the two reference seasons. Values are expressed in 10,000 Ush.

Expected yield for oilseeds are expressed in quintals/acre. Expected price for oilseeds are expressed in 10,000 Ush per Kg. Actual values of yields and price are used to construct wedges. They are expressed in the same unit of measure as the respective expectation. Wedges are constructed as the difference between actual and expected values, divided by the actual one. Hence, they are expressed as percentage deviations from the actual value.

#### *Project take-up*

Project take-up information is collected during the endline survey in both individual and group level questionnaires. The variable "VODP program take-up" is based on information on whether the farmer attended VODP2 training activities from the individual level questionnaire and on whether a learning platform was set up from the farmer group level questionnaire. The latter information is then split into the different oilseeds of interest to generate dummies for "Specific training/platform on oilseeds". We also asked the number of times farmers visited the learning platform. The variables that follow are dummies taking the value of one when the condition in the label is realized and zero otherwise.

## C Further results and robustness checks

Table C.1: First stage

	(1)	(2)
	Oilseed specific take-up	Vodp2 take-up
<b>Treat</b>	0.0778*** (0.00825)	0.471*** (0.0283)
HH head	0.0135 (0.0151)	0.241** (0.0932)
Male	-0.0148 (0.0154)	-0.241** (0.0955)
Can read	-0.00273 (0.0157)	-0.0251 (0.0576)
Can write	-0.00227 (0.0153)	0.0224 (0.0578)
No education	0.0126 (0.00961)	0.0222 (0.0550)
Primary education	0.0138** (0.00656)	0.0113 (0.0466)
Secondary education	0.0148** (0.00706)	0.0328 (0.0484)
Wealth index	-0.00213 (0.00144)	-0.00501 (0.00582)
Total land	0 (0.000100)	0.000659 (0.000569)
N. of plots cultivated	0.000878 (0.00225)	0.00527 (0.00890)
Constant	0.0934*** (0.0257)	0.246*** (0.0770)
Observations	10,172	2,543
R-squared	0.064	0.280
Controls	Yes	Yes
Crop FE	Yes	No

*Notes:* Column 1 reports the first stage of model 2b while column 2 is the first stage of model 4b. Outcome and control variables are described in Appendix B. Strata are districts. Standard errors, in parenthesis, are clustered at the level of sub-county. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

Table C.2: Program impact on oilseed adoption, robustness checks

	(1)	(2)	(3)
	Oilseed adoption	Share of land cultivated with oilseed	KLK Index
<i>PANEL A: No controls</i>			
<b>ITT</b>	0.0363*** (0.0114)	0.00631*** (0.00224)	0.0808*** (0.0258)
<b>LATE</b>	0.469*** (0.150)	0.0812*** (0.0292)	1.042*** (0.339)
<i>PANEL B: Using general VODP2 take-up</i>			
<b>LATE</b>	0.0787*** (0.0231)	0.0133*** (0.00452)	0.172*** (0.0520)
Observations	10,176	10,176	10,176
Strata FE	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes
Control mean	0.245	0.0355	0

*Note:* Panel A shows ITT and LATE as presented in model 1 and 2 but without controls (only baseline value of the dependent variable is kept), respectively. In Panel B, LATE is estimated using a general variable for VODP2 activities take-up. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table C.3: Program impact on inputs, robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Fertilizer			Pesticide			Improved seeds		KLK Index
	Use	Quantity	Expenses	Use	Quantity	Expenses	Use	Expenses	
PANEL A: No controls									
ITT	0.00649*	0.245	0.0353**	0.00736	0.0604	-0.00150	0.0258***	0.0317	0.0417**
	(0.00363)	(0.189)	(0.0135)	(0.00932)	(0.155)	(0.0120)	(0.00955)	(0.0594)	(0.0201)
LATE	0.0836*	3.205	0.461**	0.0947	0.811	-0.0201	0.333***	0.413	0.538**
	(0.0499)	(2.542)	(0.193)	(0.119)	(2.068)	(0.160)	(0.128)	(0.767)	(0.270)
PANEL B: Using general VODP2 take-up									
LATE	0.0127*	0.549	0.0751**	0.0156	0.158	-0.00264	0.0542***	0.0733	0.0892**
	(0.00772)	(0.394)	(0.0297)	(0.0199)	(0.337)	(0.0256)	(0.0205)	(0.125)	(0.0437)
Observations	10,176	9,877	9,902	10,176	9,568	9,573	10,176	9,865	10,176
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.0204	0.427	0.0215	0.0880	0.306	0.0545	0.0956	0.413	0.0174

Note: Panel A shows ITT and LATE as presented in model 1 and 2 but without controls (only baseline value of the dependent variable is kept), respectively. In Panel B, LATE is estimated using a general variable for VODP2 activities take-up. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table C.4: Program impact on labour, robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)
	Days of work on oilseeds by					KLK Index
	All	Head	Spouse	Other in the HH	Other outside the HH	
<i>PANEL A: No controls</i>						
<b>ITT</b>	0.629 (0.463)	0.310 (0.303)	0.277 (0.266)	0.0801 (0.220)	0.340*** (0.118)	0.0475* (0.0253)
<b>LATE</b>	8.159 (6.054)	4.018 (3.951)	3.586 (3.455)	1.040 (2.844)	4.420*** (1.595)	0.616* (0.335)
<i>PANEL B: Using general VODP2 take-up</i>						
<b>LATE</b>	1.416 (0.984)	0.675 (0.632)	0.656 (0.573)	0.214 (0.470)	0.732*** (0.260)	0.105* (0.0546)
Observations	9,832	9,832	9,832	9,832	9,832	9,832
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	6.210	3.593	3.515	2.344	0.617	0

*Note:* Panel A shows ITT and LATE as presented in model 1 and 2 but without controls (only baseline value of the dependent variable is kept), respectively. In Panel B, LATE is estimated using a general variable for VODP2 activities take-up. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table C.5: Program impact on market linkages, robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)
	Any sale	Seeds bought on the mkt	Harvest sold in bulk/group	Contacts with value chain actors	Sale to mkt actors	KLK Index
<i>PANEL A: No controls</i>						
<b>ITT</b>	0.0254** (0.0102)	0.00605 (0.00723)	0.00591** (0.00242)	0.00638 (0.00663)	0.000406 (0.00396)	0.0372*** (0.0140)
<b>LATE</b>	0.327** (0.129)	0.0781 (0.0906)	0.0763** (0.0320)	0.0824 (0.0848)	0.00524 (0.0507)	0.481*** (0.177)
<i>PANEL B: Using general VODP2 take-up</i>						
<b>LATE</b>	0.0544** (0.0212)	0.0139 (0.0153)	0.0122** (0.00523)	0.0124 (0.0138)	0.00147 (0.00836)	0.0785*** (0.0290)
Observations	10,176	10,176	10,176	10,176	10,176	10,176
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.130	0.0521	0.00784	0.0462	0.0263	0.0104

*Note:* Panel A shows ITT and LATE as presented in model 1 and 2 but without controls (only baseline value of the dependent variable is kept), respectively. In Panel B, LATE is estimated using a general variable for VODP2 activities take-up. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .



Table C.6: Program impact on production, robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)
	Productivity	Harvest	Harvest value	Harvest value/acre	Revenues from sale	KLK Index
<i>PANEL A: No controls</i>						
<b>ITT</b>	0.0788* (0.0412)	0.036 (0.0781)	1.914 (2.246)	4.463** (2.198)	0.621 (0.635)	0.0344 (0.0250)
<b>LATE</b>	1.018* (0.540)	0.4645 (1.00)	24.71 (28.93)	57.61* (29.64)	8.012 (8.113)	0.444 (0.324)
<i>PANEL B: Using general VODP2 take-up</i>						
<b>LATE</b>	0.177** (0.0883)	0.0805 (0.1624)	4.341 (4.802)	10.31** (4.999)	1.309 (1.347)	0.0774 (0.0536)
Observations	10,176	10,176	10,176	10,176	10,176	10,176
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.309	0.3762	13.96	16.15	3.739	0

*Note:* Panel A shows ITT and LATE as presented in model 1 and 2 but without controls (only baseline value of the dependent variable is kept), respectively. In Panel B, LATE is estimated using a general variable for VODP2 activities take-up. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table C.7: Program impact on welfare, robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)
	Total land	HH days of farm work	Rev from crop sale	HH labour income	Wealth index	KLK Index
<i>PANEL A: No controls</i>						
<b>ITT</b>	0.162 (0.536)	-4.098 (7.279)	15.61 (11.21)	-8.635 (22.40)	0.0415 (0.0654)	0.0137 (0.0267)
<i>PANEL B: Using general VODP2 take-up</i>						
<b>LATE</b>	0.346 (1.131)	-8.700 (15.18)	33.15 (23.57)	-18.36 (47.05)	0.0883 (0.138)	0.0291 (0.0565)
Observations	2,538	2,458	2,544	2,544	2,544	2,544
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
Crop FE	No	No	No	No	No	No
Control mean	7.421	218.7	131.5	139.4	0.0976	0.0197

*Note:* Panel A shows ITT and LATE as presented in model 1 and 2 but without controls (only baseline value of the dependent variable is kept), respectively. In Panel B, LATE is estimated using a general variable for VODP2 activities take-up. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table C.8: Program impact on oilseed expectations, robustness check

	(1)	(2)	(3)	(4)	(5)
	Yield expectations	Price expectations	Wedge in		KLK Index
			Yield expectations	Price expectations	
<i>PANEL A: No controls</i>					
<b>ITT</b>	-0.106 (0.846)	0.0146* (0.00749)	-0.425 (1.430)	-0.0264** (0.0121)	0.0303 (0.0214)
<b>LATE</b>	-1.117 (8.854)	0.147* (0.0764)	-4.474 (14.84)	-0.266** (0.124)	0.317 (0.220)
<i>PANEL B: Using general VODP2 take-up</i>					
<b>LATE</b>	-0.223 (1.758)	0.0310* (0.0161)	-0.784 (2.924)	-0.0566** (0.0261)	0.0653 (0.0460)
Observations	3,600	3,626	3,600	3,626	3,890
Strata FE	Yes	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes	Yes
Control mean	5.320	0.303	-3.283	0.481	-0.000531

*Note:* Panel A shows ITT and LATE as presented in model 1 and 2 but without controls (only baseline value of the dependent variable is kept), respectively. In Panel B, LATE is estimated using a general variable for VODP2 activities take-up. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table C.9: Heterogeneous effects on oilseed adoption

	Oilseed adoption	
	Levels	Wedge
Treat X Expected price	-0.0871*** (0.0276)	0.0405*** (0.0115)
Treat X Expected yield	-0.00358** (0.00165)	-0.00425 (0.00323)
Treat X Ever cultivated	0.0166 (0.0456)	0.00447 (0.0460)
Treat X Baseline adoption	0.00461 (0.0363)	-0.00324 (0.0361)
Treat X Total land	-0.000614 (0.00105)	-0.000634 (0.00105)
Treat X Wealth index	0.00289 (0.00805)	0.00148 (0.00809)
Constant	0.298*** (0.0700)	0.326*** (0.0691)
Observations	5,092	5,092
R-squared	0.106	0.104
Controls	Yes	Yes
Strata FE	Yes	Yes
Crop FE	Yes	Yes

*Note:* The table shows the heterogeneous effects of the program on oilseed adoption at the endline, by baseline values of variables. Column 1 shows the coefficients of the levels of expected price and yields at the baseline. Column 2 shows the coefficient attached to wedge in yield and price expectations at the baseline. The model also includes the main effects of the variables interacted with the treatment and of treatment itself. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Strata are districts. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table C.10: Heterogeneous effects on oilseed adoption, restricted sample

	(1)	(2)	(3)	(4)
	Oilseed adoption			
	Expected yield	Yield wedge	Expected price	Price wedge
Treat	0.0203 (0.0230)	0.00389 (0.0208)	0.0251 (0.0226)	-0.00944 (0.0203)
Het	0.00150 (0.00127)	0.00234 (0.00272)	0.00123*** (0.000176)	-0.000786*** (0.000116)
Treat X Het	-0.000448 (0.00178)	-0.00391 (0.00359)	-0.0705** (0.0295)	0.0323* (0.0171)
Constant	0.628*** (0.0641)	0.644*** (0.0626)	0.595*** (0.0615)	0.600*** (0.0612)
Observations	3,600	3,600	3,626	3,626
R-squared	0.077	0.077	0.075	0.075
Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes

*Note:* The table shows the heterogeneous effects of the program on oilseed adoption, by baseline values of variables reported in the headings, as presented in model ???. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers and the value of the outcome variable at the baseline. Outcome and control variables are described in Appendix B. Expected yield is expressed in quintal/acre, expected price in 10,000 Ush. Strata are districts. The sample is restricted to individuals with non-missing values in expectations at both baseline and endline. Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table C.11: Program impact on oilseed expectations, restricted sample

	(1)	(2)	(3)	(4)
	Expected yield	Expected price	Wedge in price	yield
ITT	-0.154 (0.684)	0.0132** (0.00630)	-0.256 (1.076)	-0.0267** (0.0102)
LATE	-1.625 (7.174)	0.137** (0.0657)	-2.699 (11.20)	-0.277** (0.109)
Observations	4,802	4,775	4,802	4,775
Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes
Crop FE	Yes	Yes	Yes	Yes
Control mean	5.320	0.303	-3.283	0.481
F-stat excl restr	65.48	61.47	65.48	61.47

*Note:* The table shows ITT and LATE as presented in model 1 and 2 in the text, respectively. Baseline controls include respondent characteristics (head of household, male, can read, can write, schooling), household characteristics (wealth index, total cultivated land, n. of plots cultivated), dummies for the assigned service providers. Outcome and control variables are described in Appendix B. Expected yield is expressed in quintal/acre, expected price in 10,000 Ush. The sample includes farmers with non-missing expectations at the endline only. Strata are districts. The Robust standard errors clustered at the level of sub-county in parentheses. Asterisks denote statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .