

The Role of Information in Agricultural Technology Adoption: Experimental Evidence from Rice Farmers in Uganda

Bjorn Van Campenhout, Wilberforce Walukano and Piet Van Asten

June 29, 2016

Introduction

Abstract

Previous research identified information inefficiencies as a major constraint to sustainable crop intensification among rice farmers in Eastern Uganda. The fact that some farmers report not using certain inputs or techniques because they are not aware of them while others report they are aware of them but are not using them suggests information gaps at two levels. First, farmers may lack knowledge about the existence or use a particular input or technology. Second, a farmer may be unable to correctly evaluate the returns to using the technology. In this study we therefore try out two different information treatments at the individual level. In a first intervention, we show farmers the recommended practices and inputs in rice farming, using an agricultural extension video message. In a second intervention, we point out the returns to investment in a series of simple calculations that also consider the longer run. The study uses a 2 by 2 factorial design with randomization over matched blocks of four farmers. This document provides some background information, presents methods we will use, outlines hypotheses which will be tested, defines outcome variables to be used and specifications we plan to estimate. As such, it will provide a useful reference in evaluating the final results of the study (Humphreys, Sanchez de la Sierra, and van der Windt, 2013).

Motivation

This study is motivated by the fact that, in the near future, food production will need to increase substantially to cater for ever more people with more calorie intensive diets. In many parts of Africa, agricultural yields, defined as crop produced per measure of land, are very low. To increase yields, much is expected from sustainable intensification, similar to the green revolution in Asia. There, modern inputs and technologies, such as inorganic fertilizers, improved seeds, the use of herbicides and pesticides, and a range of improved cultivation practices were able to increase productivity in a relatively short period of time (Mueller et al., 2012). Especially in Uganda, a country that combines one of the highest growth rates in Africa with a predominantly agriculture based economy, sustainable crop intensification is important for both food security and poverty reduction.

The present study is a follow-up on an exploratory study on sustainable crop intensification among rice farmers in eastern Uganda. Due to the government's efforts to promote rice, it has become an important cash crop for people living in eastern Uganda, where extensive swamps and high rain-fall create conducive agro-climatic conditions. With cassava and matooke¹ affected by pests and diseases, rice is also becoming important for food security. According to the Uganda Census of Agriculture, rice production has increased from about 40,000 MT to about 200,000 MT between 2000 and 2009. More than 2/3 of this is produced in the East.

Past research suggested that, similar to other crops, there is a substantial yield gap for rice in Uganda (Van Campenhout, Bizimungu, and Birungi, 2016, Figure 3.1 on p.5). We also find that, in observational data, the use of fertilizer, pesticides and recommended practices such as timely transplanting and proper water management is associated with significantly higher yields. While some of the farmers did not use these intensification methods because they are not aware to them, others reported they knew them but did not practice them, often because they thought they were not profitable. This suggests knowledge gaps at two levels:

First, there are farmers that simply **lack technical knowledge**. That is, they are not aware of the existence of the particular improved input or technique or they do not have the knowledge on how to use or implement it. Increasing yields through changing inputs and practices requires that knowledge about these inputs and techniques reaches the farmers (Jack, 2011). In

¹Matooke is a starchy cooking banana.

rural and remote areas, where infrastructure is lacking, education levels are low and farmers have been using traditional methods and inputs for ages, generating this technical knowledge poses a challenge. In addition, as information is a public good, private entities tend to under-invest in it. As a result, most agricultural information is provided through a public agricultural extension system, where extension agents visit farmers in their communities. However, monitoring difficulties and a lack of accountability often result in limited scale, sustainability and impact of these extension systems (Aker, 2011).

Second, farmers **lack of knowledge about profitability** of an input or technology may affect adoption². Investments in agricultural inputs or technologies requires the farmer to compare costs today to a stream of future income. However, farmers may not have precise information about the returns to investment, and as such base their decision to invest on *perceived* return. There are different reasons to suspect perceived returns to crop intensification investment are lower than actual return. It may be that farmers are not sophisticated about the yield response to inputs or techniques. In addition, performing inter-temporal cost-benefit analysis may prove challenging for poorly educated farmers (Cole et al., 2013). Behavioral economics also suggests human beings are present biased and thus tend to underestimate returns that are in a distant future (Frederick, Loewenstein, and O'Donoghue, 2002). In addition, human beings tend to overestimate rare events in judgment tasks (Gonzalez and Wu, 1999). If farmers are risk averse and associate more risk to modern inputs and techniques, this will again reduce the expected return to investment. Additional biases such as pessimism bias, where negative outcomes are considered more likely, and safety-first behavior, where poor farmers avoid even the smallest uncertainties to ensure survival, are likely to further reduce perceived returns to sustainable crop intensification.

A similar mismatch between perceived and real returns is encountered in investment decisions in education. Also here, parents need to make a decision on additional schooling (incurring both direct costs and indirect costs related to the opportunity cost of the child that would otherwise enter the labor market) based on the perceived returns to this additional schooling.

²Profitability of for instance investment in fertilizer is unlikely to be homogeneous across rice farmers in the area. Indeed, the existence of transaction costs at virtually all levels affects profitability, leading some households to self select out of markets (de Janvry, Fafchamps, and Sadoulet, 1991). This implies that some intensification investments will remain unprofitable for at least some farmers even after relaxing information constraints.

It has been found perceived returns are often too low, and simply providing information on the actual returns increases schooling outcomes (Jensen, 2010).

We have chosen to conduct this study among smallholder rice farmer in Eastern Uganda. The area is located South of lake Kyoga, a large shallow lake that provides good conditions for rice growing. To the east is Kenya while lake Victoria is located on the South. Rice yields average at about 2MT/ha and slightly higher if we restrict attention to the main rice growing season (second season of 2013). Still, there is a significant productivity gap when compared to potential yields as obtained in research stations (about 5 MT/ha). In addition, within the area, yields vary substantially, and the distribution is skewed to the right, with many farmers having lower than average yields and a few farmers having very high yields. For instance, while median yields are only about 1.7 MT/ha, the 10 percent farmers with highest yields attain 3.6 MT/ha (Van Campenhout, Bizimungu, and Birungi, 2016).

Research Questions

The research investigates (a) whether information inefficiencies are responsible for low levels of agricultural technology adoption and subsequent yields; and (b) the effectiveness of farm level interventions that aim to reduce information inefficiencies at both the technical level and/or the level of expected profitability. In particular, we will estimate the effect of increasing knowledge of recommended practices on outcomes such as technology adoption and yields. In a second experiment, we will estimate the effect of providing information on the profitability of different investment on outcomes such as technology adoption and yields. We also want to look at interaction between these two information gaps. To make sure that our interventions affect outcomes through relaxing knowledge constraints, we will also measure some intermediate variables. In particular, at the time of the interventions, we will try to get an idea about the technical knowledge of farmers, and test if knowledge changes in response to the treatment. Similarly, we will try to get an idea of the perceived returns to crop intensification investments and also test if our intervention changes these perceptions.

We will label the first type of information constraints **technological information barriers (TI)**. To quantify the importance of this constraint for technology adoption and sustainable crop intensification, we will design a simple intervention that relaxes this constraint. In particular, we will

show simple extension video messages that make farmers aware of existing inputs, technologies and best practices. This will be done at the individual level using Android tablet computers, as it has been argued that Information and Communication Technologies (ICTs) may provide a cost effective way to address these information-related barriers to technology adoption (Aker, 2011). For instance, The Grameen Foundation uses smart-phones in Uganda to provide extension information through a Community Knowledge Worker (Van Campenhout, 2016). Cole and Fernando (2013) evaluate a mobile-phone based technology that allows farmers to call a hot-line, ask questions and receive responses from agricultural scientists and local extension workers.

The second type of information constraint will be referred to as **returns to investment information barriers (RI)**. Also for this information constraint, we will develop a simple information treatment, where farmers are taken through the calculations of returns to different investments. For instance, we will show the farmer what it would cost to apply fertilizer to his or her field and what the expected return would be. We will also pay some attention to effects that go beyond one agricultural season, covering areas such as reinvesting profits, recycling improved seeds or the longer run benefits of proper soil management. Information treatments have been found to be effective in many other settings where perceptions are biased. For instance, Alexis Grigorieff (2016) find that biased beliefs about the number of immigrants in a country can be reduced through simple information treatments. Financial literacy has also been shown to affect uptake of insurance among smallholder farmers (Cole et al., 2013).

We are interested in the effect of the information that is provided in the video, not in the effect of being shown a video. Therefore, we need to account for the effect that comes from merely being shown a video. This can be done by including a placebo video. For instance, Bernard et al. (2015) use music videos as a placebo in their experiment of the effect of showing videos of successful farmers. However, for our case where we have rather specific information in each video, showing a completely unrelated video such as a music video may be unsatisfactory. One could still argue that the effect comes from being shown a video on rice farming instead of from the technical knowledge or returns to investment information that is provided in the videos³. Therefore, we decided to produce a placebo video that resembles the

³Note that the factorial design does allow answering this question, by comparing the pure TI treatment to the pure RI treatment. However, for this test, we can only use half

treatments as closely as possible. In particular, we will produce a video that provides information on post harvest handling⁴. While we choose the topic of the video such that it is unlikely to affect outcomes that are of interest in this study, this can not be entirely excluded. For instance, it may be that farmers that get information on post harvest handling anticipate higher returns and also start changing practices during growing, which would affect yields. We therefore decided to administer the placebo treatment to all participants.

In this research, we define effectiveness in terms of changes in a range of outcome variables. The ultimate outcome of interest is the level of welfare of smallholder farmers. We will thus look at a number of outcomes that are correlated to welfare. Some of these outcomes may be quantitative and absolute in nature (such as consumption expenditure per capita or per adult equivalent) while others may be subjective and relative (for instance if a farmer considers him or herself better or worse off than other farmers). However, we will also investigate how these outcomes are affected by looking at the impact on intermediate outcomes. For instance, welfare may be affected directly through higher yields. Alternatively, households may get higher prices for their rice in the market as a result of higher quality due to the use of improved practices. We will thus also estimate the impact on farm gate prices and on yields. This process can be refined and more immediate outcomes can be investigated. For instance, we will gather information on pests and diseases to find out how the interventions affect yields. A detailed list of what variables will be included in the study is given below.

Research Strategy

Sampling

Sampling Frame

We study smallholder rice farmers in eastern Uganda. The eligible sample for this study are rice farmers that will be planting rice in the second season of 2016. Planting for this season starts around August and can go on up to October. Between June and August 2014, we collected detailed socio-economic data from about 400 small-holder rice farmers in the East of Uganda. In

of the sample, and the test is likely to be under-powered.

⁴This video will also provide useful information to participants, reducing the ethical objection to working with a control group that one group does not benefit.

particular, we sampled from three districts: Tororo, Butaleja and Bugiri. Sampling of households was done with the assistance of the Uganda Bureau of Statistics (UBOS).

Table 1 reports some statistics of the households in our sample. For instance, we find that households consists are on average of about 7.45 members, which is considerably larger than average household size in Uganda. The table also shows that 93 percent of the households are headed by a male household head. Average age of the household head is 42 years. Table 2 further shows that about 60 percent of the household heads can read and write and 75 percent of households have access to a mobile phone. In general, farmers in the area have little land (less than 2 hectares). Farmers have to walk about 23 minutes to get to their land.

For rice in particular, we find average yields in the second season of 2013 was about 2 MT/ha of milled rice equivalent. The average area allocated to rice production is about 0.85 hectare. Most rice farmers have some experience. The average farmer has been growing rice for more about 11 years. Only about 23 percent of farmers use fertilizer on their rice fields. More than three quarters of rice is sold in the market, underscoring the importance of rice for the livelihoods of farmers in the East. Only a small fraction of the harvest is saved for seed.

Statistical Power

Our interventions are expected to reduce knowledge gaps related to technology and profitability, which in turn is expected to increase the adoption of sustainable crop intensification methods such as fertilizer and pesticides among rice farmers. We therefore look at the different in mean yields between farmers that use a particular input or technology and those that do not in our baseline data to get an idea about what effect size to expect. To do so, we restrict ourselves to farm households that reported growing rice in the second season of 2013.

We find that average yields are about 2.11 MT/ha for farmers that report not using fertilizers. Farmers that use fertilizer report yields of about 3MT/ha, which is an increase of about 42 percent. The pooled standard deviation is about 1.29 MT/ha. To detect such an effect, we would need about 26 observations in each treatment arm (single sided, 80 percent power and alpha level of 0.05). For pesticide use, we find that farmers that do not report using pesticides get yields of around 2.16 MT/ha, while those that do

use fertilizer obtain about 2.84, corresponding to an increase of about 31 percent. We would need about 45 observations in each treatment arm to detect such an effect. For recommended practices, we find that farmers that maintain water depth of 10-25 cm during cultivation to effectively control weeds have 41 percent higher yields, and we would need about 40 observations to identify such an effect with 80 percent power. Finally, farmers that plant in rows have average yields of about 2.71MT/ha, while those who do not plant in rows attain about 2.17MT/ha. To find this effect, we would need about 68 observations in each treatment arm to detect this effect.

The above are the effects of implementing a particular technology. However, our information treatment is unlikely to encourage all farmers to start using improved inputs or technologies. Therefore, we expect effects to be smaller and settle for a 20 % average increase in yields. To identify such an effect, we need about 110 observations in each treatment arm. We therefore propose to run an experiment that involves about 220 observations in a 2x2 factorial design. In such a design, about 110 households will receive the TI treatment. About 110 households will receive the RI treatment. This will be done in such a way that there are 55 households that receive both TI and RI and 55 households that do not receive RI nor TI. Note that, as explained above, all 220 farmers will (also) get the placebo treatment.

Assignment to Treatment

Pairwise-matching has been found to be superior to re-randomization in small samples. Matching on covariates can increase balance on these covariates, and increase the efficiency and power of hypotheses tests (Bruhn and McKenzie, 2009). In addition, King et al. (2007) point out an additional advantage: if a unit drops out of the survey, its paired observations can also be dropped without compromising overall balance. This is different in conventional randomized experiments, where if one observation drops out, it can no longer be guaranteed that treatment and control groups are on average balanced.

We will exploit data from the baseline survey to match blocs of four individuals along a range of observable characteristics (Greevy et al., 2004). Therefore, we wrote an algorithm that clusters observations on the basis of euclidean distance. In particular, we randomly select a farmer (i) from the sample of eligible farmers and pair this farmer to the most similar farmer in terms of a range of characteristics (such as age if household head, land acreage under cultivation,...). This is done by minimizing the square root of the sum

of squared standardized differences of the measures for these characteristics. At the same time, we want to maximize the distance between the farmers to reduce potential spillover effects⁵:

$$\min_{i,j} \left(\sqrt{(age_i - age_j)^2 + \dots + (ha_i - ha_j)^2 - (lat_i - lat_j)^2 - (long_i - long_j)^2} \right) \quad (1)$$

After household (j) is identified, it is given the same bloc number as observation i and the second best matching household is determined. When this is found, it is also given the same block number as household i and j . Thereafter, the third best matching household is determined, which is also given the same block number. Now we have matched the first four households. The block is then removed from the sample and the entire process starts again. This is repeated until the desired number of blocs are formed. We have matched 252 farmers in 63 blocs.

In general, one attempts to obtain balance on variables that are thought to be strongly correlated with outcomes of interest. We have matched using the following general household characteristics: household size, age of household head, sex of household head, area under rice cultivation, distance to nearest input provider, access to credit and whether the household has received training or extension services in rice growing in the last 5 years. It is important to note that we also included key outcome variables itself such as consumption per capita and rice productivity as characteristics to match on. Finally, as mentioned above, we maximize distance between households as measured by GPS coordinates. While these covariates explain about 16 percent of the variation in baseline (log) productivity and about 21 percent of variation in baseline (log) expenditure per capita. The R program that does the sampling can be found at <https://is.gd/BA2LjW>, while the actual sampling list with the allocation of treatments can be found at <https://is.gd/iRQ6xm>.

⁵The exact implementation differs slightly. For instance, we relate the difference in latitude and longitude to the maximum differences within the sample. Also, we maximize distance between each farmer and all the other farmers within a bloc. The algorithm is implemented as an R function and is available at <https://is.gd/YyFEtG>.

Fieldwork

Instruments

Apart from the baseline data that was collected as part of the PASIC project, we will be using 2 instruments in the field. First, there are the actual interventions that consist of information treatments in the form of short (about 5 minutes) videos and will be shown to individual farmers in the field. The videos will be embedded in a short questionnaire that asks some questions, mainly for validation or to test hypotheses about the impact pathways. The second instrument will be a standard survey to collect end-line information on a range of outcome variables. The instruments will be described in detail below.

The first instruments will be developed from scratch. We will produce one video for the **technical information treatment (TI)** and one video for the **returns to investment treatment (RI)**. In the latter treatment, we will go through some of the calculations together with the farmer, and so this video will have a large component that tries to improve financial literacy. To make the videos, we had extensive interviews with farmers and experts on rice growing in the region. From these interviews we have distilled the most important steps which have been converted into scripts. These criteria for the steps were that they should have a large effect on productivity, and few farmers should be using them. The choice of what interventions will feature in the videos was also be informed by the relationships found in the baseline data.

For the TI treatment, the video will focus on 3 skills (timeliness, water management and fertilizer use)⁶:

- It is important that you plan well ahead.
- Before sowing, prepare bunds in the rice field.
- bunds should be about 60 cm high and 50 cm width.
- plough the field, this is easier when the ground is moist

⁶Note that initially, we also included pesticides as a key intensification input. However, discussions with experts taught us that pesticides are not really necessary. In addition, we wanted to avoid promoting the use of toxic substances, some of which are imported from countries that lack proper health regulations.

- after ploughing, submerge the field
- at this time, start the nursery, sowing X kg of clean seed per acre
- start paddling and leveling of the submerged field
- 14 days after sowing, reduce water to at most 1 inch and makes sure all weeds are gone
- apply NPK (or DAP) at 25 kg per acre while transplanting
- transplant 2 seedlings per hill, 1.5 inches deep 20 cm from between each hill, preferably in rows.
- then keep field moist for 14 days
- add water up to 2 inches and remove weeds
- apply urea 25kg/acre
- keep the water *in* the field for 14 days - if water levels drop due to eg evaporation, add water to maintain optimal levels of 2 inches
- five weeks after apply second time of urea after weeding, again 25 kg per acre. Again, keep water in the field for 14 days.
- increase water levels up to 3 inches, this depth should be maintained up to maturing.
- One week before harvest, gradually reduce water to zero by harvesting day.

In this movie, we will avoid alluding to the results of these efforts. In particular, we will avoid contrasting yields from farmers who use fertilizer to farmers who do not. We will also avoid showing how pesticides increases plant health. In short, we want the video to respond to the “how” question, while avoiding the “why” questions.

For the RI treatment, the video will be much less of a how-to, but more of a video that points out facts relevant for cost benefit analysis. For instance, it will highlight the cost of inputs and the effect of inputs and methods on yields. It will teach farmers some basic financial literacy and encourage farmers to take a long run perspective when making decisions. We will make

sure farmers do not get to see images that provide clues on how to do things. The video will focus on fertilizer, timeliness and water management, which mirror those mentioned in the how-to video:

1. basic facts about building blocs of inter-temporal decision making:
 - (a) returns
 - Timeliness: transplanting in time increases yields by 3.5 bags per acre.
 - Proper water management reduces yields by 3 bags per acre
 - fertilizer use on average increases yields by 3.6 bags per acre
 - (b) costs
 - cost of fertilizer
 - cost of labor
2. cost benefit analysis: show that the additional cost of inputs and technologies is much smaller than the additional revenue of inputs and technologies - illustrate return on investment (1000 sh of fertilizer - 1 hour of labor)
3. taking a longer run view (explaining sunk costs and recurrent revenues):
 - mention that proper water management (without becoming too specific and show how to manage water) has long run effects on soil since it conserves top soil.
 - illustrate how profits from fertilizer can be reinvested and generate progressively bigger stream of income (start small)
 - show planning and managing the farm as a commercial entity transforms lives...

This video should provide answers to the “why” questions while avoiding the “how” question.

Finally, we will also produce a video to be shown to all participants, including the control. This placebo intervention mixes techniques and issues related to profitability. As such, it can contain parts that show how to do things but also explain why it is important, as well as the ultimate goal of increasing quality of end product and getting a higher price for rice. The video will go over the following aspects of post harvest handling:

- paddy must be put on a tarpaulin
- paddy should be spread out at thickness of about 5 cm
- drying paddy should be shuffled every 30 minutes to allow equal exposure to sun
- drying paddy should not be left in direct sunlight for more than 3 hours, dry slowly to avoid broken rice
- dry like this for 3 to 4 days
- don't sell rice immediately after harvest, but wait 3 months. If you are in urgent need for money, sell only part and make sure you know the going price in the market.

The videos will be shot by a professional videographer that also has experience in communications.

The use of information treatments as the interventions has some obvious advantages. First, the use of a pre-recorded video results in a standardized treatment, and all subjects receive exactly the same treatment. While one may argue that providing the information through trainers may be more effective, as the trainer may adapt the message to eg. the education level of the recipient, this may also lead to subtle differences in the message given. The videos will also be administered at the individual level. Again, one may argue that providing the information at a more aggregate level, such as to cooperatives, may be more cost-effective. However, it will be very difficult to control group dynamics, and thus providing information to groups may again lead to heterogeneous treatments. We also use video to reduce spill-over effects. For instance, an alternative to a video would be to provide posters or brochures that explain the use of fertilizer and pesticides and to inform farmers about profitability. This may actually be more effective, as farmers can keep these materials and get back to them at different points in time. The video will be shown only once and farmers may forget some the recommendations over time. However, providing printed material can more easily be passed on to neighbors and relatives, potentially contaminating other treatment or control groups. Illiterate farmers also are likely to benefit more from videos than from written material. Finally, the provision of a relatively hands-off information treatment (instead of for instance providing inputs) was also chosen because we want to evaluate an intervention that is

cheap and easy to scale up in a setting that is more realistic than the typical experimental field trials used in the agronomy studies.

We will ask a series of questions immediately after farmers have seen the video's. This is to see if farmers learned something from the interventions and also to investigate impact pathways. The following questions have been tested in the field and will be asked to each farmer immediately after the video(s) were shown:

- Are you growing rice?
- Are you using fertilizer?
 - depending on answer above:
 - * Suppose you are farming one acre of rice as you normally would, that is, (not) using fertilizer. In a bad year, with adverse weather conditions, how much would you get?
 - * And in a good year with favorable weather conditions?
 - * Now suppose you are farming one acre of rice as you normally would, except that you are now (not) using fertilizer. In a bad year, with adverse weather conditions, how much would you get?
 - * And in a good year with favorable weather conditions?
- Do you normally transplant between 14 and 21 days after sowing?
 - depending on answer above:
 - * Suppose you are farming one acre of rice as you normally would, that is, (not) transplanting between 14 and 21 days. In a bad year, with adverse weather conditions, how much would you get?
 - * And in a good year with favorable weather conditions?
 - * Now suppose you are farming one acre of rice as you normally would, except that you are now (not) transplanting between 14 and 21 days. In a bad year, with adverse weather conditions, how much would you get?
 - * And in a good year with favorable weather conditions?
- How long should rice be in the sun each when drying each day? (multiple choice)

- 3 hours
- half day
- entire day
- When is the best time to transplant? (multiple choice)
 - after 15 days
 - after 25 days
 - after 35 days
- When drying your rice, you should: (multiple choice)
 - spread as much as possible on tarpaulin
 - spread 5 cm (2 inches) thick on tarpaulin
 - in small heaps on tarpaulin
- How often should you apply UREA to rice? (multiple choice)
 - once
 - twice
 - 3 times

The video's will be screened on Samsung Galaxy Tab 2 tablet computers. We will develop an Open Data Kit ODK application which has ID data pre-loaded, such that the correct video(s) are cued without the intervention of the enumerator. The ODK app will also be used to capture responses to the control question. At this time, we will also check and if necessary update some contact data.

The second instrument, the end-line survey that will be administered at around January 2017, will be based on an existing survey that was used to collect the baseline data. However, this was a very lengthy survey that included many socio-economic characteristics. For the end-line, we will concentrate on collecting information on the outcome variables that are listed below. The survey will be implemented as CAPI (Computer Assisted Personal Interviewing) using Open Data Kit (ODK) on Samsung Galaxy Tab 2.

Empirical Analysis

Variables

The following variables are all relevant to our study and will be collected during end-line and used in the analysis. Some of these variables are also used in the orthogonality tests.

- direct yield related
 - total rice production
 - area of paddy planted
 - area of paddy planted as share of total area
 - log(rice yields): estimated kg of milled rice produced over the course of an agricultural season divided by ares in hectares. We will trim upper and lower 5 % of observations with absolute cut offs at 100 and 10000 kg/hectare.
- sales, which may also indicate difference in quality
 - amount sold in kg
 - price at which is sold
 - timing of sales
 - to whom sold
- Well-being and food security
 - log(consumption per capita)
 - rice consumption
 - crop portfolio
- Intensification technologies
 - use of fertilizer
 - use of pesticides
 - use of recommended practices such as row planting.

Balancing Checks

In general, matching doesn't necessarily guarantee the balance of any particular covariate. However, by design, we expect to find balance at baseline on the variables included in the matching procedure. In addition, Bruhn and McKenzie (2009) find that in small samples of less than 300 observations, and with very persistent outcomes, matching on relevant baseline variables achieves more balance in follow-up outcomes. We will thus make a clear distinction between balance tests for variables that are included in the matching algorithm and those that are not. We will run simple regressions with the outcome for which we want to check balance on a treatment indicator and a set of bloc dummies (see equation 3 below). We also run joint orthogonality tests, where we regress the treatment indicator on all variables we want to jointly test for balance (in addition to the bloc dummies). We will then judge balance by looking at the F-statistic.

The results for a set of variables that are also included in the matching procedure are in table 1. The second column, titled *mean*, provides sample means for each of the variables. For example, we find that among the 252 sampled rice farmers, the average household size is just under 8 persons and that average age of the household head is 42 years. Below each mean, we provide standard errors in brackets. The third column compares balance between the control group and the group that will receive the video that provides information on existing intensification inputs and technologies (TI treatment). In particular, it shows the difference in the average outcome at baseline between farmers that will be exposed to the treatment and those that will not. Given our factorial design, those that will be exposed to the treatment include farmers that will receive only the TI treatment (63) but also those that will be shown both the TI video and the RI video (also 63). Similarly, the control for this case are the 63 farmers that will not receive any treatment at all (control) and those that will receive only the RI treatment (again 63 farmers). In a similar fashion, the fourth column compares average outcomes for farmers that will be exposed to the RI treatment to those that will not. The final column tests the balance of the crossed treatment (TI+RI). It compares the average of the 63 farmers that get both the TI treatment and the RI treatment to the 63 farmers that do not get any treatment at all.

As expected, we find no systematic violation of the orthogonality conditions for variables that were included in the matching procedure. For in-

stance, we see that the average household consists of about 7.45 individuals. Farmers that are earmarked to receive the TI treatment are on average about 0.2 individuals larger. However, the difference is not statistically different from zero. On average, the age of the household head is about 42 years. Farmers that will receive the TI treatment are on average 2 years younger. Here we do find a significant difference between treatment and control at the 10 percent significance level. We also find some imbalance in the distance to input provider. Here, rice farmers that will receive the TI treatment are significantly closer to an input dealer, as well as rice farmers that will receive both treatments.

For the joint orthogonality tests, we run three different regression. One regression has as a dependent variable an indicator that is one if the household gets the TI treatment and zero otherwise. A second regression has as the dependent variable an indicator that is one if the household will receive the RI treatment. Both of these regressions will use all 252 observations. Finally, a third regression will use an indicator as the dependent variable that is one if the household will get both TI and RI treatments and zero if the household gets none of these treatments. This regression is based on only 126 observations. Doing so for the TI treatment, only age of the household head shows up significant at the 10 percent level. The F-statistic is 0.148. For the RI treatment none of the variables are significant and the F-statistic is 0.123. For the cross treatment, distance to input provider is significantly negative at the 5 percent levels. The F-statistic of 0.193. For none of the three regressions, joint orthogonality is rejected.

We also test a range of other variables which were not used for matching. The choice of variables was based on what variables other researchers in similar studies use in their orthogonality tests. In particular, we looked at balance tables in studies that investigate the adoption of yield improving methods and technologies using RCTs. These studies include Duflo, Kremer, and Robinson (2011), Karlan et al. (2014), Ashraf, Giné, and Karlan (2009), Bulte et al. (2014) and Matsumoto (2014). In addition, we also add some variables that are relevant in the context of a study on sustainable crop intensification among rice farmers. The results are reported in table 2. As was to be expected, balance is violated more frequently now. For instance, we find that in the sample of farmers that will receive the TI treatment, farmers are less likely to be able to read and write. In this group, farmers are also setting aside a larger share of their harvest for seed. They also have significantly more experience in rice growing as proxied by the years they

	mean	TI	RI	TI+RI
household size	7.45 (2.81)	0.20 (0.232)	0.13 (0.232)	0.33 (0.332)
age of head	41.96 (12.90)	-1.96+ (1.078)	1.66 (1.081)	-0.30 (1.313)
log rice prod	7.48 (0.71)	0.06 (0.057)	-0.02 (0.057)	0.04 (0.091)
male head	0.93 (0.25)	0.01 (0.014)	0.01 (0.014)	0.02 (0.016)
area of rice	0.85 (0.58)	0.05 (0.050)	0.02 (0.050)	0.07 (0.076)
log welfare per capita	7.67 (0.47)	0.02 (0.039)	0.03 (0.039)	0.05 (0.063)
distance to input provider	6.27 (6.61)	-0.98+ (0.538)	-0.65 (0.540)	-1.63* (0.762)
credit (/1000)	115.36 (373.37)	-4.76 (26.116)	-38.25 (25.968)	-43.02 (40.560)
extension	0.15 (0.35)	0.01 (0.022)	0.02 (0.022)	0.03 (0.039)
nobs	252	252	252	126

Table 1: Balance table for variables used in matching

	mean	TI	RI	TI+RI
can read and write	0.59 (0.49)	-0.12* (0.055)	0.04 (0.056)	-0.08 (0.083)
total ha under cultivation	1.67 (1.40)	0.01 (0.164)	0.07 (0.164)	0.08 (0.266)
average time to reach parcel (min)	23.06 (21.88)	2.30 (2.874)	4.29 (2.862)	6.89+ (4.107)
experience in rice growing (yrs)	11.35 (8.64)	1.77+ (0.956)	-0.15 (0.965)	1.62 (1.516)
use of fertilizer on rice	0.23 (0.42)	0.06 (0.049)	-0.02 (0.049)	0.03 (0.075)
share of harvest kept for seed	0.05 (0.06)	0.02** (0.008)	-0.01 (0.008)	0.01 (0.010)
share of harvest sold	0.76 (0.22)	-0.01 (0.029)	0.02 (0.029)	0.01 (0.040)
has mobile phone	0.75 (0.43)	-0.02 (0.056)	-0.07 (0.056)	-0.10 (0.074)
distance to nearest market	4.94 (4.72)	1.37* (0.595)	-0.84 (0.600)	0.53 (0.875)
nobs	252	252	252	126

Table 2: Orthogonality tests

have been growing rice. Finally, farmers that will receive the information treatment live on average 1.37 km farther away from the market.

We also run joint orthogonality tests, producing F-statistic between 0.359 and 0.236, and so joint orthogonality can not be rejected. The reported balance tests are based on baseline data. In the final study, the same tables will be produced with end-line data. The same variables will be used and the same specifications run. The R code that is used to generate these tables is publicly available in the version control system (<https://is.gd/BQRPN9>)

Treatment Effects

Intent to Treat

We will run three different specifications to identify the treatment effects (β). First, note that in each bloc $b = \{b_1, \dots, b_{62}\}$, we have four treatments that were randomly assigned. For the main treatments, each bloc always has two treated and two control observations $t = \{c_1, c_2, r_1, r_2\}$. We start by simply estimating the average treatment effect of an outcome variable (y):

$$y_{t,b} = \alpha + \beta T_{t=r,b} + \varepsilon_{t,b} \quad (2)$$

where T is an indicator function that is one if $t = \{r_1, r_2\}$ and zero otherwise. This specification is for reference only.

Equation 2 does not account for the matched randomization which we outlined in the section on sampling above. Not accounting for the method of randomization may result in overly conservative standard errors and a significant reduction in power (Bruhn and McKenzie, 2009)⁷. Therefore, a second specification includes fixed effects for the blocs:

$$y_{t,b} = \alpha + \delta_b + \beta T_{t=r,b} + \varepsilon_{t,b} \quad (3)$$

which exploits variation within each block. Defining bloc means of the outcome variable $\bar{y}_b = \frac{1}{4} \sum y_{t,b}$

$$(y_{t,b} - \bar{y}_b) = \beta \left(T_{t=r,b} - \frac{1}{2} \right) + (\varepsilon_{t,b} - \bar{\varepsilon}_{t,b}) \quad (4)$$

Finally, for outcome variables for which we also collected baseline data, such as rice yields, we also run difference-in-difference regressions, where we now have observations in two rounds $s = \{z, e\}$:

$$y_{s,t,b} = \alpha + \delta_b + R_{s=e} + T_{t=r,b} + \beta T_{t=r,b} \cdot R_{s=e} + \varepsilon_{r,t,b} \quad (5)$$

$$(y_{s=e,t=r,b} - y_{s=z,t=r,b}) - (y_{s=e,t=c,b} - y_{s=z,t=c,b}) = \beta + \varepsilon_{r,t,b} \quad (6)$$

Wherever specification allows, we will rely on randomization inference to consistently estimate standard errors in small samples (Barrios et al., 2012).

⁷One may argue these type II errors not such a problem. However, (Bruhn and McKenzie, 2009) also find that for a substantial part of random allocations, there is an increase in type I errors. We will therefore only consider results from specifications that control for our matching design as evidence.

Treatment on the Treated

Given the nature of our experiment, the issue of compliance that often creates a discrepancy between the intention to treat and the actual treatment of the treated is very low. That is, we do not expect a significant part of the selected farmers to refuse to look at the video. Still, we can think about the experiment also as an encouragement design, where the information treatments are designed to encourage farmers to intensify. As such, it may be possible that some of the farmers that view a particular video do not actually engage in crop intensification and so we will not expect a change in outcomes such as yields. Therefore, we will run equation 3 again, but replace the treatment indicator ($T_{t=r,b}$) by an indicator for the actual investment in crop intensification methods, which will then be instrumented by ($T_{t=r,b}$). To do so, we will thus include a series of questions in the end-line study that enables us to find out if the farmer uses pesticides and fertilizer, and whether he or she is following recommended practices with respect to water management and transplanting according to what was featured in the videos.

Heterogeneous and Spillover Effects

Information treatments are likely to be affected by education levels and training. We will therefore also investigate heterogeneous effects related to education level. In addition, we will look at interactions with previous exposure to agricultural extension. However, we did not consider heterogeneous effects in our power calculations, so we do not expected much statistical power.

Randomization is at the individual level and thus spillover effects, leading to a downward bias in our estimates, are a concern. While our matching procedure was designed to minimize spillover effects within blocs by maximizing distance between matched farmers, it is still possible that for instance control farmers from another bloc resides close to a treated farmer and gets exposed to the treatment⁸. Therefore, following Cole and Fernando (2016), we will also run specifications that control for the fraction of a respondent's social network that was also part of the study and the fractions that were assigned to the different treatments.

⁸Although the chance that this control farmer interacts with the treated farmer from the other bloc should be smaller since he or she is less similar.

Standard Error Adjustments

While our main outcome variables are ultimately household welfare and rice yields, we will also estimate the impact of our interventions on a range of intermediate variables. The fact that we have many such variables may lead to the so-called “look elsewhere” effect, where one is bound to find significant effects simply due to the sheer number of parameters. Therefore, some form of multiple-inference correction is in order. In general, there are two ways in which to avoid false positives that result of multiple hypothesis testing. One can either reduce the number of hypothesis, or one can make the statistical test stricter by for instance reducing the significance threshold (such as the Bonferroni adjustment). We will address false positive arising from multiple hypothesis testing using both ways.

First of all, we will use the groupings presented in the section that lists the variables to create indices (directly related to yield, sales, welfare, crowding in other intensification methods,...). At the most basic level, each of the indices is a weighted mean of the several standardized outcomes within each group. In particular, for each variable within each group, we make sure positive direction always means better, otherwise we switch sign. We then demean the outcome and standardize by scaling by the control group standard deviation. We then create weighted averages for the outcomes in each group at the household level, using as weights the inverse of the co-variance matrix of the transformed outcomes within the group. This is done for each of the groups. The resulting variables can then be used to assess the impact of the particular intervention using the specifications outlined above.

However, we may be interested in identifying differential effects within each of the groups. For example, we may want to differentiate between the effect on rice sales immediately after the harvest and rice sales during the lean season. We will therefore also use Family Wise Error Rate Control. In particular, we will use the free step-down re-sampling method of Westfall and Young (1993).

The matched design of our experiment is likely to introduce dependence among outcome variables within blocs, and so we need to account for clustering when estimating an equation like 2. As the typical approach to deal with this, using cluster robust standard errors, is known to be biased in small samples, we will use randomization inference (RI). Instead of relying on a theoretical distribution, RI involves comparing the test statistic to the distribution of the test statistic under each possible allocation of treatments.

In particular, within each bloc of four matched farmers, we compute the outcome (eg. proportion of correct answers) for all 6 possible permutations of two treatments (as the four treatment arms in the factorial design correspond to two pure treatments). For 63 blocs, this leads to a total of 6^6 , so instead of actually computing all combinations, we base inference on 10000 random samples from this distribution. Finally, we will also drop outcomes from our analysis for which 95 percent of observations are the same value. This is done to reduce the influence of outcomes with limited variation.

Research Team

The research will be led by Bjorn Van Campenhout (b.vancampenhout@cgiar.org), with assistance of Piet Van Asten (p.vanasten@cgiar.org). Research assistance will be provided by Wilberforce Walukano (W.Walukano@cgiar.org) and Marc Charles Wanume (wcharli@gmail.com).

Deliverables and Calendar

- research plan and pre-analysis plan registration (AEARCTR-0001312)
- IRB approval - under review.
- Interventions: 3 videos of about 5 minutes each by end of July 2016.
- Fieldwork for interventions – 126 to be shown TI video, 126 to be shown RI video, all to be shown placebo, last week of July 2016, first week of August 2016.
- report of intervention: October 2016
- End-line survey: Jan-Feb 2017
- Data analysis, report writing and dissemination: July-August 2017

References

Aker, J. C. 2011. "Dial "A" for agriculture: a review of information and communication technologies for agricultural extension in developing countries." *Agricultural Economics* 42 (6): 631–647.

- Alexis Grigorieff, D. U., Christopher Roth. 2016. *Does Information Change Attitudes Towards Immigrants? Evidence from Survey Experiments*. Tech. rep.
- Ashraf, N., X. Giné, and D. Karlan. 2009. “Finding missing markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in Kenya.” *American Journal of Agricultural Economics* 91 (4): 973–990.
- Barrios, T., R. Diamond, G. W. Imbens, and M. KolesÅr. 2012. “Clustering, Spatial Correlations, and Randomization Inference.” *Journal of the American Statistical Association* 107 (498): 578–591.
- Bernard, T., S. Dercon, K. Orkin, and A. Seyoum Taffesse. 2015. “Will Video Kill the Radio Star? Assessing the Potential of Targeted Exposure to Role Models through Video.” *The World Bank Economic Review* 29 (suppl 1): S226–S237.
- Bruhn, M. and D. McKenzie. 2009. “In Pursuit of Balance: Randomization in Practice in Development Field Experiments.” *American Economic Journal: Applied Economics* 1 (4): 200–232.
- Bulte, E., G. Beekman, S. Di Falco, J. Hella, and P. Lei. 2014. “Behavioral responses and the impact of new agricultural technologies: Evidence from a double-blind field experiment in tanzania.” *American Journal of Agricultural Economics* 96 (3): 813–830.
- Cole, S., X. Gine, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery. 2013. “Barriers to Household Risk Management: Evidence from India.” *American Economic Journal: Applied Economics* 5 (1): 104–35.
- Cole, S. A. and A. N. Fernando. 2013. *The Value of Advice: Evidence from Mobile Phone-Based Agricultural Extension*. 2013 AAAE Fourth International Conference, September 22-25, 2013, Hammamet, Tunisia 160520, African Association of Agricultural Economists (AAAE).
- . 2016. *‘Mobile’izing Agricultural Advice: Technology Adoption, Diffusion and Sustainability*. Working Paper 13-047, Harvard Business School.

- de Janvry, A., M. Fafchamps, and E. Sadoulet. 1991. "Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained." *Economic Journal* 101 (409): 1400–417.
- Duflo, E., M. Kremer, and J. Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *The American Economic Review* 101 (6): 2350–2390.
- Frederick, S., G. Loewenstein, and T. O'Donoghue. 2002. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature* 40 (2): 351–401.
- Gonzalez, R. and G. Wu. 1999. "On the Shape of the Probability Weighting Function." *Cognitive Psychology* 38 (1): 129 – 166.
- Greevy, R., B. Lu, J. H. Silber, and P. Rosenbaum. 2004. "Optimal multivariate matching before randomization." *Biostatistics* 5 (2): 263–275.
- Humphreys, M., R. Sanchez de la Sierra, and P. van der Windt. 2013. "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration." *Political Analysis* 21 (1): 1–20.
- Jack, K. 2011. *Constraints on the adoption of agricultural technologies in developing countries*. Tech. rep., J-PAL and CEGA.
- Jensen, R. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *The Quarterly Journal of Economics* 125 (2): 515–548.
- Karlan, D., R. Osei, I. Osei-Akoto, C. Udry et al. 2014. "Agricultural Decisions after Relaxing Credit and Risk Constraints." *The Quarterly Journal of Economics* 129 (2): 597–652.
- King, G., E. Gakidou, N. Ravishankar, R. T. Moore, J. Lakin, M. Vargas, M. M. Téllez-Rojo, J. E. H. Ávila, M. H. Ávila, and H. H. Llamas. 2007. "A "Politically Robust" Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program." *Journal of Policy Analysis and Management* 26: 479–506.
- Matsumoto, T. 2014. "Disseminating new farming practices among small scale farmers: An experimental intervention in Uganda." *Journal of the Japanese and International Economies* 33: 43–74.

- Mueller, N. D., J. S. Gerber, M. Johnston, D. K. Ray, N. Ramankutty, and J. A. Foley. 2012. "Closing yield gaps through nutrient and water management." *Nature* 490 (7419): 254–257.
- Van Campenhout, B. 2016. "There's an App for that? The Impact of Community Knowledge Workers in Uganda." *Information, Communication & Society* .
- Van Campenhout, B., E. Bizimungu, and D. Birungi. 2016. *Risk and Sustainable Crop Intensification - The Case of Smallholder Rice and Potato Farmers in Uganda*. Discussion Paper 1521, International Food Policy Research Institute.
- Westfall, P. and S. Young. 1993. *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*. A Wiley-Interscience publication. Wiley.