SMS-extension and Farmer Behavior: Lessons from Six RCTs in East Africa*

Raissa Fabregas, Michael Kremer, Matthew Lowes, Robert On, Giulia Zane†

September 1, 2019

For most recent version click here.

Abstract

This paper provides new causal evidence on the effects of SMS-based agricultural extension programs on farmer behavior. We present results from six RCTs conducted with farmers in Kenya and Rwanda. All six programs encouraged farmers residing in areas with acidic soils to experiment with agricultural lime, an input that can reduce soil acidity and increase yields. Four programs also encouraged farmers to experiment with certain types of fertilizers. Programs varied in their design, informational content and target populations. To interpret the findings, we use meta-analytic techniques to combine the results. Our odds ratio estimates for the effects of the programs on purchases of agricultural lime is 1.21 (95% CI: 1.13, 1.28) and for purchases of fertilizer is 1.13 (95% CI: 0.85, 1.51). While we find evidence of increases in knowledge, we do not find that increases in adoption are sustained over the following season. We do not find compelling evidence to suggest that providing additional details about local soil characteristics, and access to a call center, significantly increased the impact of simple text messages. However, using experimental variation from two different programs, we find that repeating the same messages had a statistically significant impact on the adoption of inputs. Overall, the effect sizes are modest, but relative to the low cost of text-messages, these programs can be cost-effective.

*Funding from ATAI, the Dioraphte Foundation, Unorthodox Philanthropy, 3ie, and an anonymous donor made this study possible. We thank our partner organizations: the Kenya Agriculture and Livestock Research Organization (KALRO), and in particular to Dr. Martins Odendo, One Acre Fund (OAF), Precision Agriculture for Development (PAD), and Innovations for Poverty Action - Kenya (IPA-K) for all of their support during implementation and data collection. We are also grateful to OAF, the Mumias Sugar Company, and German Agro Action (Welthungerhilfe) for sharing soil data and specially to Dr. Javier Castellanos for providing the agronomic recommendations to PAD and Dr. David Guerena for the agronomic recommendations for OAF. We thank Lillian Alexander, Carolina Corral, Tomoko Harigaya, Charles Misiati, Chrispinus Musungu, Cara Myers, Alex Nawar, Carol Nekesa, William Wanjala Oduor, Violet Omenyo, Victor Perez, and Megan Sheahan for excellent project and research support. This project benefited from feedback from Rema Hanna, Rohini Pande, Kashi Kafle, Jessica Zhu and participants at AAEA, MWIDC and NEUDC. Disclosure: Michael Kremer is a board member of PAD. Giulia Zane's post-doctoral fellowship was funded by PAD. Fabregas, Kremer, and On do not receive any financial compensation from the organization but the organization partly funded this research. On worked as a volunteer for OAF at the time of the experiment.

†Fabregas: UT Austin, LBJ School of Public Affairs, rfabregas@utexas.edu; Kremer: Harvard Economics Department and NBER, mkremer@fas.harvard.edu; Lowes: One Acre Fund, matt.lowes@oneacrefund.org; On: Chan Zuckerberg Initiative, roberton@gmail.com; Zane: Inter-American Development Bank, giuliaz@iadb.org.

1 Introduction

The widespread adoption of mobile phones in developing countries potentially allows governments and other organizations, to deliver information at scale, targeting individuals' specific circumstances, in a timely manner and at a very low cost. Hundreds of different public and private sector digital information initiatives have been deployed in developing and emerging economies (Aker, 2017; Aker and Mbiti, 2010). While only a fraction have been evaluated, there is a growing interest in understanding whether these approaches are effective in changing individual behavior.

There is evidence that programs that rely on short message services (SMS), one of the cheapest ways to deliver information, can improve educational outcomes (Ksoll et al., 2014; Aker et al., 2012; Cunha et al., 2017), financial behaviors (Karlan et al., 2012, 2016), civil servant performance (Dustan et al., 2018) and increase civic and political engagement (Aker et al., 2017; Buntaine et al., 2018; George et al., 2018). In other sectors, such as agriculture and health, the empirical evidence has been described as mixed, with evidence of some positive and significant and some statistically insignificant effects (Aker, 2017). Having limited and heterogeneous evidence complicates the decision of whether to expand or continue these types of programs. If program effectiveness is very sensitive to specific features of its design, the identity of the implementing organization, or the local context, it might be difficult to draw broader lessons about the impact of these programs.

In this paper, we present new evidence on the impacts of six different SMS-based agricultural extension programs on farmer behavior. To systematically summarize the results and learn about the extent of heterogeneity that exists across studies we aggregate the evidence using meta-analytic techniques. The programs were implemented in Rwanda and Kenya by three different organizations: the Kenya Agriculture and Livestock Research Organization (KALRO), One Acre Fund (OAF) and Precision Agriculture for Development (PAD) jointly with Innovations for Poverty Action (IPA). All programs encouraged farmers to experiment with specific agricultural inputs, in particular agricultural lime (an input used to reduce soil acidity) and some types of chemical fertilizers. However, there was program variation in terms of farmer targeting, informational content, complementary services offered and experimental design. This set-up, therefore, helps us study a common situation that arises for development projects: different implementation agencies may have similar goals (e.g. increase experimentation with agricultural lime in acidic soils) and tools (i.e. SMS) but they go about implementing their programs in different ways.

Agricultural extension has been one of the primary approaches used to encourage technology adoption among farmers. However, traditional in-person extension efforts have been criticized for being deficient, expensive and affected by governance issues (Anderson and Feder, 2007). Therefore, there has been a growing interest in using phones to deliver this information directly to farmers. This paper focuses on delivery through text-messages. One reason for this is that the most common type of phone used in developing countries is still the basic device with only call and texting capabilities, and delivery through SMS is still one of the cheapest ways to reach people at scale (World Bank, 2012). From the perspective of carriers, the marginal cost of delivering text-messages is close to zero and even with positive pricing bulk SMS is extremely cheap.[1] On the other hand, illiteracy, message complexity or the cognitive cost of sorting through messages could limit the impacts of conveying information over text. More broadly, informational interventions will only be effective if behavior change is not completely constrained by other market imperfections (Aker et al., 2016).

The evidence we present complements previous work focused on the role of information and communication technologies (ICT) on agricultural development. Existing review articles, such as Aker et al. (2016) and Nakasone et al. (2014), provide an in-depth overview of different approaches and services related to agriculture. These reviews have called for additional evidence on the effectiveness of ICT-based extension services (Nakasone et al., 2014), but have also noted that while these systems appear to increase knowledge, they "have little to no impact on agricultural practices, production, or farm-gate prices"(Aker, 2017).

[1] In 2018, we document that some services in Kenya charge less that $0.006 per SMS and we found services in India that varied anywhere from $0.006 to $0.0004 depending on the number of messages bought.

A complication in drawing conclusions from the existing literature is that the effect sizes required for very cheap programs to be cost-effective are very small. Therefore, to have sufficient power to detect these effects, sample sizes have to be fairly large.[2] Opportunities to experiment at scale might be limited and costly. One of the contributions of this paper is to provide new evidence based on experiments that are large enough to detect small effects. Combining the evidence through a meta-analysis further increases statistical power.

Depending on the study we have up to two types of data on farmer behavior: survey data, and data either from redemption of input discount coupons that were distributed to both treatment and control farmers (IPA/PAD, and KALRO), or from the administrative records of OAF. Using administrative data as our primary outcome mitigates concerns around reporting bias in farmer self-reported behavior, whereas the use of survey data allows us to explore additional outcomes such as knowledge increases and impacts on the use of other inputs. An additional advantage of using administrative data, is that it helps us overcome some of the financial difficulties of collecting individual information with large sample sizes. This highlights another issue that might constraint researchers' ability to measure the impacts of these types of programs: collecting other indicators, such as physically measuring yield increases, would be prohibitively expensive at the required scale.

A few other evaluations of phone-based extension have measured changes in farmer behavior, though with some exceptions they have mostly relied on self-reported data. A study of a SMS-extension program offered to sugar cane farmers in Kenya found positive yield impacts in one trial but no effects on a second trial with a different sample (Casaburi et al., 2014). Larochelle et al. (2019) study a SMS-based program for potato farmers in Ecuador. The authors find that the intervention increased knowledge and self-reported adoption of integrated soil management practices. Van Campenhout et al. (2018) evaluate the effect of an ICT-based extension program for maize farmers in Uganda. The authors find impacts from extension videos, but no additional effects from adding SMS reminders and voice recorded messages. Perhaps suggesting that informational needs had already been addressed by the video intervention. Cole and Fer-

[2] A similar point is made by Lewis and Rao (2015) about measuring the return to advertisement.

nando (2016) evaluate an advisory service targeted at cotton farmers in India and find increases in self-reported adoption of recommended agricultural inputs for cotton cultivation. However, unlike the programs in this paper, that service was delivered through a hotline and not through text-messages. Two other evaluations (Fafchamps and Minten, 2012; Camacho and Conover, 2010) have used SMS to reach farmers but have mainly focused on the effects of crop price and weather information rather than extension services. The service evaluated by Fafchamps and Minten (2012), however, also contained some crop advisory messages. The authors do not find any evidence indicative of farmers changing behavior or practices. A different weather and price information service implemented in Colombia, found improvements on knowledge but not on prices farmers' received or profits (Camacho and Conover, 2010).

Using estimates from our six experiments, we find that the combined odds ratio for the effects of being enrolled in the SMS extension programs on lime purchases is 1.21 (95% CI 1.13 to 1.28, N=6) and the effect for purchasing fertilizers was 1.13 (95% CI: 0.85 to 1.51, N=4). For the lime results, we fail to reject the null of homogeneous effects across programs, and the percentage of variability, which measures the share of variability not explained by sampling error, is low though imprecise (I^2=17%, 95% CI 0% to 62%). We reject the null of homogeneity and find more substantial heterogeneity for the fertilizer results, though again the results are imprecise (I^2=68%, 95% CI 6% to 89%).

There might be a number of different mechanisms through which text messages might affect farmer behavior. On the one hand, messages may simply increase awareness about the existence of inputs or increase knowledge about how to use them. On the other hand, the programs might remind farmers to use the inputs, or could help them aggregate different signals about the adequacy of these technologies. While we cannot fully disentangle whether these effects are driven by increases in knowledge, persuasion or reminders, we highlight three findings. First, we find that the programs increased knowledge about soil acidity and lime. The combined odds ratio for identifying lime as a way to address soil acidity is 1.57 (95% CI 1.40 to 1.76, N=4). Second, the lime adoption effects are not sustained over the following season. Third, we cannot reject that impacts are the same for those farmers who did not know much about lime at baseline and those who did. This might suggest that the effects are not consistent with a simple knowledge transmission story. Moreover, we do not find evidence that these programs led to a reduction of farmers' use of other inputs or practices. Suggesting that, at least in these setting, displacement effects or limited cognitive bandwidth to focus only on certain technologies might not be a primary concern.

Using experimental variation from different treatment arms within selected studies, we draw some lessons about different features that could strengthen these programs. First, we do not find that differences in message framing made much of a difference on the adoption of lime. However, message repetition is effective at increasing purchases: there are significant effects from receiving a second message on the likelihood of purchasing lime. Other add-ons, such as talking to an extension officer through the phone or providing additional information to farmers based on local soil data, had no additional effect relative the general SMS messages.

This paper is organized as follows: section 2 describes the context that we study and provides some agricultural background. Section 3 presents the design of each program and their evaluations. Section 4 discusses the empirical strategies. Section 5 provides the results. Section 6 provides some discussion on some of the lessons that we can draw from the programs. We present cost-effectiveness estimates in section 7. We conclude in section 8.

2 Maize Farming and Input Use in East Africa

The projects we study targeted maize farmers across Rwanda and western Kenya between 2014 and 2017 (see Figure 1 for a map).[3] In these regions, as in many parts of Sub-Saharan Africa, smallholder yields have remained low, partly because of issues of soil degradation and nutrient depletion, soil acidity, and low adoption of productivity-enhancing technologies.

High soil acidity, corresponding to pH levels below 5.5, can dramatically reduce crop yields by limiting nutrient availability to the plants (The et al., 2006; Tisdale et al., 1990; Brady and

[3] All programs targeted a specific agricultural season. In both countries maize is farmed twice a year. In Kenya, the primary agricultural, the long rain season, takes place from March until August and a secondary agricultural season, the short rains season, takes place from September until December. In Rwanda, the main agricultural season takes place from September to January and the secondary season takes place from March to August.

Weil, 2004) and it is associated with aluminum and manganese toxicities, which can inhibit plant development (NAAIAP, 2014).[4] It is estimated that over one third of sub-Saharan African soils are acidic (Pauw, 1994). The application of agricultural lime to the soil is one of the cheapest and most widely recommended methods to increase soil pH. Experimental plots conducted in Kenya suggest that lime application can increase maize yields by 5-75% (Kisinyo et al., 2015; Gudu et al., 2005; OAF, 2014).[5] While agricultural lime is cheap, it is bulky, so one limitation to broadcasting large quantities of lime in a field is that it might be difficult for smallholder farmers to transport it and store it. Therefore, a second approach has been to recommend farmers to micro-dose it to each planting hole (one or two soda bottletops per hole). This approach entails a lower dosage but requires re-application each season. OAF reports that micro-dosing lime in experimental plots increased yields by at least 14% (OAF, 2014). With the exception of the KALRO program, all services recommended experimenting with micro-dosing.

Currently, several public agencies and NGOs in Africa have advocated for the use of lime. Yet, farmers have limited knowledge about lime. For instance, in Kenya, at baseline only 25% of farmers participating in the second IPA/PAD program knew that lime could be used to reduce soil acidity and only 9% of them reported having ever used it. In addition, whether to use lime and the optimal quantity to apply depends on soil chemistry.[6] Few smallholder farmers in this region conduct soil chemistry tests in their own farms as they are not easily accessible and relatively expensive.[7] To better target farmers, IPA/PAD and OAF aggregated field-level soil information to the area-level to make predictions about their soil acidity. KALRO simply

[4] In particular phosphorus becomes less available to the plant. This can also imply that the use of fertilizers is less efficient on these soils.

[5] These estimates reflect results for trials with and without combining lime with other inputs, particularly fertilizers containing nitrogen and phosphorous. For instance, OAF's experimental plots suggest that broadcasting lime evenly over maize fields before the planting season begins application, in combination to the standard fertilizer package they recommend, increased yields by 25% (OAF, 2014).

[6] The optimal level of pH is between 5.5 and 7 (NAAIAP, 2014), applying too much agricultural lime is not only inefficient, but it can lead to alkaline soils which cannot sustain crops (Kiplagat et al., 2014).

[7] A wet soil test for all nutrients through public agency is at least \$11, a soil test through private company can reach over \$20. This does not account for other costs such as transportation of samples, materials, etc.

recommended farmers to test their own soils.[8]

In addition to lime, four programs provided information about less commonly known types of fertilizers, with the objective of encouraging farmers to experiment with them. A large body of work suggests that chemical fertilizer can substantially raise agricultural yields (Evenson and Gollin, 2003), and previous research in the region suggests that certain fertilizers are profitable if used in the right quantity (Duflo et al., 2008). There is a range of different fertilizers available in this area which differ in chemical composition, soil and crop suitability, and price. Most farmers in Western Kenya have used specific types of phosphorus and/or nitrogen-based chemical fertilizers for maize, such as diammonium phosphate (DAP). However, fewer farmers have experimented with other options, particularly top-dressing fertilizers, such as calcium ammonium nitrate (CAN) and urea. Which fertilizer is most profitable, can depending on local conditions. For instance, per dollar spent, urea contains more nitrogen per unit than CAN. However, its effectiveness depends on rain availability since most of the nutrients can be lost to evaporation if not dissolved in water (Overdahl et al., 2017). Therefore, there might be gains from better targeting information to farmers' local conditions. Other fertilizers, such as Mavuno, were specifically blended to target micronutrient deficiencies in Western Kenya.

3 The SMS Extension Services and the Experimental Designs

Table 1 summarizes the six programs. We show characteristics of the organizations, messages and the agricultural season and location where the experiments took place. In Table 2 we describe the features of each evaluation. A full description of each program and its corresponding evaluation design can be found in appendix A.

The programs we study were implemented between 2015 and 2017 in different regions of

[8] Using area-level soil information rather than field-level information to predict soil acidity, could be a potentially scalable strategy that could be used to provide better targeted recommendations to farmers (relative to no information). There are a number of different efforts seeking to generate and compile localized soil data with the goal of improving management of soils.For instance, a number of projects have been launched to gather soil data, e.g. africasoils.net, soilmap.org, soilgrids.org, etc. In a separate project, using soil data, we document that using area-level means rather than global means reduced the mean squared error of the prediction by 12% for pH (Fabregas et al., 2017b).Since there might also local variation in pH (Tjernstrom et al., 2015), farmers in the IPA/PAD programs were advised to experiment in a small portion of their land with the inputs.

Kenya and Rwanda. Three different organizations implemented these programs: KALRO, a Kenyan public agency with the mandate to promote agricultural research and dissemination in the country; PAD, a non-profit organization that supports the provision of phone-based customized agricultural information services to smallholder farmers, jointly with IPA a research non-profit organization; and OAF, a non-profit social enterprise that provides training and agricultural inputs on credit to smallholder farmers. In particular, OAF works with groups of 8 to 11 farmers, who sign input contracts before the agricultural season starts. The inputs are then delivered by OAF right before the beginning of the following planting season. Farmers repay their input loans at any time during the growing season.

Farmers who participated in the KALRO SMS program were recruited through a village census conducted by IPA. Those enrolled in the first IPA/PAD program (IPA/PAD1-K) were found through existing databases.[9] Farmers enrolled in the second IPA/PAD program (IPA/PAD2-K) were recruited through agricultural supply dealers (agrodealers) who invited their clients to participate. All OAF SMS programs (OAF1-K, OAF2-K, OAF3-R) targeted previous OAF clients.

The SMS programs were popular among farmers. All of the farmers who were invited to participate in the KALRO program opted-in. In the IPA/PAD1-K and IPA/PAD2-K, 95% and 99.5% of the farmers invited agreed to received the messages respectively. There was no formal opt-in process for the OAF messages, since they were sent as part of regular OAF activities. Two programs (OAF1-K and IPA/PAD2-K) offered the additional option to talk to an agricultural field officer over the phone. However, only 1-8% of farmers offered this add-on took advantage of it.[10]

There was some variation in the number of messages sent by each program, ranging from 5 to 28. The message content also varied. For instance, in addition to sending lime and fertilizer recommendations, the KALRO and IPA/PAD1-K programs also provided suggestions around

[9] Farmers in Busia were registered in an existing IPA database for a large-scale farming project, those in Kakamega were part of a database of a large agrobusiness in the region.

[10] Farmers were invited to 'flash' the organizations (i.e. dial the number and hang up, so that they wouldn't pay for the call). The extension officer would then all them back.Therefore, farmers did not have to incur a cost to use this additional service.

Table 1: Program Characteristics

	KALRO (1)	**IPA/PAD1-K (2)**	**IPA/PAD2-K (3)**	**OAF1-K (4)**	**OAF2-K (5)**	**OAF3-R (6)**
Org. Type	Public	NGOs	NGOs	Social Enterprise	Social Enterprise	Social Enterprise
Location	Kakamega and Siaya (Kenya)	Busia and Kakamega (Kenya)	Busia, Bungoma, Kakamega & Siaya (Kenya)	Busia and Kakamega (Kenya)	Bungoma, Busia, Kakamega and Vihiga (Kenya)	Western, Eastern, Southern (Rwanda)
Agricultural Season	SR 2015	SR 2016 & LR 2017	LR 2017	SR 2016	SR 2017	Main Season 2017
Recruitment	Farmers drawn from village census	Former NGO and contract farming participants	Clients of agrodealer	OAF clients in Long Rains 2016	OAF clients in Long Rains 2016	OAF clients in 2017
Eligibility	Phone owner, farmed during past year, in charge of farming	Planted maize in 2016, reside in program area	Clients of agrodealers	OAF clients in 2015	OAF clients in 2016	OAF clients in 2017
Message Content	Lime, fertilizer, seeds, field management	Lime, fertilizer, field management	Lime and fertilizer	Lime	Lime and fertilizer	Lime
Number of Messages	20 total (2 acidity/lime; 5 fertilizer)	24-28 total (7-9 acidity/lime; 4-9 fertilizer)	13 total (6 acidity/lime; 4 fertilizer)	6 total (6 acidity/lime;0 fertilizer)	1-10 total (1-5 acidity/lime; 1-5 fertilizer)	1-4 total (1-4 acidity/lime; 0 fertilizer)
Lime recommended?	All (if acidic)	0.81	0.76	All	All	All
Fertilizer recommended?	DAP, NPK, CAN, Mavuno	Urea	Urea	No	CAN	No
Used Local Soil Data?	No	Yes	Yes	Yes	Yes	Yes
Additional Services?	No	No	Phone-call	OAF Services & Call-center	OAF Services & Call-center	OAF Services & Call-center
Message Repetition	No	Yes	Yes	Yes	Yes	Yes
Opt-in	1	0.95	0.95	-	-	-
Previous lime use^b	0.06	0.12	0.09	-	-	0.06
Previous fert. use^b	0.84	0.86	0.84	0.95	0.93	0.95
Female^b	0.65	0.37	0.34	0.64	0.69	-
Primary School^b	0.53	0.60	0.72	-	-	-

Notes: SR denotes Short Rain Season (August-January) and LR Long Rain Season (March-August) ^{b} denotes data for control group at baseline. - denotes that data is unavailable. Lime recommended indicates whether all farmers received messages recommending positive amounts of lime, or the fraction that did. Fertilizer recommended whether fertilizer messages were sent, and if yes, the types of fertilizer. Opt-in indicates the fraction of farmers who when invited agreed to received SMS.

other agricultural practices. The messages sent by IPA/PAD2-K and OAF were mostly focused on lime, with the exception of one treatment arm within the OAF2-K program which also sent messages about a specific type of top-dressing fertilizer.

The programs implemented by OAF and IPA/PAD used data from soil tests to construct their messages (details in appendix C). Based on this data, IPA/PAD did not recommend positive amounts of lime to all farmers. Those in areas where the median pH was over 5.5 were recommended not to apply lime (corresponding to 18% of sample in PAD/IPA1-K and 23% in PAD/IPA2-K). Therefore, we define our outcome variables as 'following the recommendations' rather than just purchasing the input.[11]

Using baseline data for each experiment, we can look at some demographic characteristics. About two-thirds of participants are females, except in the IPA/PAD samples, where the proportion of women is reversed. Where we have survey information over 50% of farmers report completing primary school. For both the OAF3-R and KALRO samples, 6-12% of respondents report having used or purchased agricultural lime in a previous season. Most farmers had used some type of chemical fertilizer in the previous agricultural season.

From the six experiments, five are individual randomized trials with randomization at the farmer level (KALRO, IPA/PAD1-K, IPA/PAD2-K,OAF1-K,OAF2-K) and one is a clustered randomized trial, with randomization at the farmer group level (OAF3-K). In the OAF3-K trial, farmer groups were randomized into treatment and control, and then half of the treatment groups were randomized to be fully treated or partially treated. The objective of this design was to be measure spillovers. For our main analysis we exclude within-group controls. In section 5 we discuss the extent of these spillovers.

The sample size for each experiment ranged from 800 farmers (KALRO) to 110,400 farmers (OAF3-R). Except for KALRO, which only had one treatment arm, all experiments had several treatment arms. In this paper we analyze data pooling all treatment arms together for each experiment. This increases power and simplifies the analysis and discussion. However, we also

[11] E.g. coded as one if the farmer used lime and lime was recommended or if the farmer did not use lime and lime was not recommended.

Table 2: Research Design

	KALRO (1)	IPA/PAD1-K (2)	IPA/PAD2-K (3)	OAF1-K (4)	OAF2-K (5)	OAF3-K (6)
Unit of randomization	Individual	Individual	Individual	Individual	Individual	Cluster (farmer group)
Sample Size	834	1,897	5,890	4,884	32,572	110,400
Treatment Arms (#)	1	2	2	2	2+	2+
Treatment Arms	**1.SMS**	**1.General SMS** and **2.Specific SMS**: sent additional information about local acidity level, input prices and quantities.	**1.SMS**, **2.SMS + Call**: also received call by field officer, **3. SMS + Call offer**: offer to receive phone call	**1.Broad SMS**, **2.Detailed SMS** additional info on degree of soil acidity, lime quantity, cost, and predicted yield increase.	**1.Lime only**: Lime SMS. Cross-randomized message framing, repetitions and frequency. **2.Lime+CAN**: additional messages encouraging also to buy extra CAN.	**1.Same message**: All farmers in a group got same SMS. Cross-randomized message framing, repetitions and frequency. **2.Different messages**:All farmers in group got different SMS. Cross-randomize message framing, repetitions and frequency.
Admin Outcome	Coupon (paper)	Coupon (digital)	Coupon (digital)	OAF admin	OAF admin	OAF admin
Coupon Value	50% discount lime,50% discount fertilizer	Choice 10 Kg lime or soap (first season); 15% discount lime (second season); 30% discount CAN, Urea	15% discount lime; 15% discount fertilizer	-	-	-
Baseline Survey	Yes	Yes (phone)	Yes (phone)	No	No	No
Endline Survey	Yes	Yes (phone)	Yes (phone)	Yes (phone)	No	No
Data Collection	SR 2015 (survey)& Coupon (LR 2016)	SR 2016 (coupon) & LR 2017 (survey)	LR 2017 (coupon) & LR 2017 - SR 2017 (survey)	SR 2016 (purchases) & LR 2017 (survey)	SR 2017 (purchases)	June 2017 (purchases)

Note: SR and LR denote the Short and the Long Rain agricultural season in Kenya, respectively. Treatment arms (#) denotes the number of treatment arms, for OAF '+' indicates that there were cross randomizations in these samples for the number of messages (1-5), frequency sent, and framing (7 possibilities).

provide tables with results for each individual treatment arm in Appendix G and highlight some lessons from these experimental variations in section 6.

To measure changes in farmer behavior we observe an indicator of input acquisition for all programs. These 'administrative' sources might be less prone to reporting bias, since farmers have to put resources on the line. The administrative data for OAF programs consists of agricultural input orders. The OAF programs messaged previous farmer clients before farmers had signed up loan contracts for that agricultural season. Only 60-76% of farmers who received OAF text messages signed-up again to receive OAF loans. While we do not find evidence of a differential likelihood of requesting a loan by treatment status (appendix E, table E7), we take a conservative approach and define our outcome variable as lime purchased through OAF, without conditioning on whether farmers were in fact OAF clients at the time of the experiment.

For KALRO and IPA/PAD we use data from coupons redeemed at local agricultural supply shops. The coupons, provided to all farmers in treatment and control groups, were devised as a way to collect information on input choices and reduce concerns about enumerator demand effects. Agricultural shops were incentivized to keep records of coupon redemption and farmer input choices. Farmers in the KALRO sample received two paper coupons redeemable for a 50% discount for lime and fertilizer. IPA/PAD coupons were sent via SMS, and either provided discounts or a choice between inputs or a good of similar value (to address liquidity constraints). Finally, we also have self-reported endline data for four programs. An in-person endline survey was collected with KALRO farmers. Endline surveys with questions about input use were conducted over the phone with IPA/PAD farmers and with a randomly selected sub-sample of OAF2-K farmers. Since all these different sources of data were collected at different times we have survey and/or administrative data for two agricultural seasons for each program.

3.1 Validity of the Experimental Designs

Balance tables for equality of means for each experiment and reports on survey attrition can be found in appendix E. For OAF programs, for which we did not collect baseline data, we use OAF administrative data from previous seasons. Overall, the samples appear to be balanced across most characteristics. We reject the null of joint significance in all cases, except for OAF1-K were we find small differences at baseline: those in the treatment arms were more likely to purchase onion seeds and additional CAN fertilizer, and to receive a repayment incentive the previous year. We control for these variables in our main specifications, but the results are robust to their exclusion. Appendix D contains a list of controls used in all regression specifications. The follow-up rate in all experiments ranged from 75% (IPA/PAD1-K) to 91% (KALRO) (only a random one third of farmers in the experiment were attempted to be interviewed for OAF1-K). We do not find any evidence of differential attrition by treatment status (table E7).

4 Empirical Strategy

Main outcomes are 'following lime' and 'following fertilizer' recommendations. For programs for which we collected survey data, we can also measure changes in agricultural knowledge. In all cases we estimate intention-to-treat effects.[12] We pool all treatment arms for the IPA/PAD and OAF programs. The general equation we estimate for each program is:

$$y_i = \alpha + \beta Treatment_i + X_i \nu + \gamma_w + \epsilon_i, \tag{1}$$

where y_i is the outcome measure for farmer i. $Treatment_i$ denotes a dummy variable indicating treatment, X_i is a vector of controls for farmer specific characteristics, γ_w controls for area fixed effects and ϵ_i, is the error term. The coefficient β estimates the difference between treatment and control. For binary outcomes, we estimate equation 1 with a logistic regression model and report the coefficient β in terms of odds ratios (OR) for the probability of acquiring the input. In the appendix we also show results for linear probability specifications.

In order to improve precision and address some small baseline imbalances we control for the strata used in each randomization, demographic characteristics, farming practices, previous

[12]It is possible that some farmers might not have received the messages.For instance, anecdotally, some farmers reported that only those with access to the main network in the area could receive messages during the IPA/PAD1-K. Some farmers in OAF3-R did not receive messages because they did no own a phone and had listed someone else's number.

input use, location fixed effects, and for the survey data we include enumerator fixed effects. Finally, since the randomization was at the group level for the OAF3-R experiment the errors terms are clustered at that level.

To synthesize the evidence across these various experiments and present a weighted average of study estimates, we conduct a meta-analysis. We use a random effects model, which assumes that true effects in each study are normally distributed.[13] Formally, for each experiment we observe an estimated treatment effect:

$$\hat{T}_j = \theta_j + e_j \tag{2}$$

where θ_j is the true effect for study j, and e_j is the within-study error, where $e_j \sim N(0, \sigma_j)$, and σ_j is the sampling variation in estimating θ_j. We further assume that $\theta_j = \mu + \delta_j$, and $\delta_j \sim N(0, \tau^2)$, where τ^2, the between-study variance, is estimated by the DerSimonan and Laird method (DerSimonian and Laird, 1986). Our estimate of μ is:

$$\hat{\mu} = \frac{\sum_{j=1}^{s} w_j T_j}{\sum_{j=1}^{s} w_j} \tag{3}$$

Where w_j are study specific weights given by the inverse of the variance. In this case,

$$w_j = \frac{1}{(\hat{\tau}^2 + \hat{\sigma}_j^2)} \tag{4}$$

In addition to τ^2, we report two other measures of heterogeneity across programs, Cochran's Q statistic to test the null hypothesis of homogenous effects across studies, and, since this test has low power when the number of studies is small (Higgins et al., 2008), Higgin's and Thompson's I^2, the percentage of variability not explained by sampling error (Higgins et al., 2003; Higgins and Thompson, 2002). We complement this analysis by running heterogeneity specifications pooling all datasets together and estimating a single model (as in equation 1) with strata controls and

[13]In comparison, in the context of meta-analysis, a fixed-effect model would assume that all the studies share a common true effect.The random effects model is more plausible in our context, since true effect sizes might vary by population, the specifics of the intervention or contextual factors.

experiment dummies.

5 Results

5.1 Awareness and Knowledge about Lime

Text-messages increased the proportion of farmers who knew lime was a remedy for soil acidity (knowledge). Rows 1 and 2 in table 3 show that the treatment effects as an odds ratio for farmers having heard of lime (awareness) is 1.21 (95% CI 0.94 to 1.57). There is substantial heterogeneity in this result, the p-value of the Q statistic is 0.02 and I^2=69.7% (95% CI 13% to 89%). The odds ratio for knowing that lime can reduce soil acidity is 1.57 (95% CI 1.40 to 1.76). We cannot reject the null of homogeneous treatment effects on knowledge. The p-value of the Q statistic is 0.626 and I^2=0 (95% CI 0% to 85%). Figure 2 shows a forest plot for these results. Table F1 in appendix F shows effects for each program separately.

5.2 Following Recommendations on Agricultural Lime

We now examine one of the key behaviors that all programs were expected to affect: following the lime recommendations. For the PAD/IPA programs, we code the recommendation as being followed if the farmer used lime and lime was recommended or if the farmer did not use lime and lime was not recommended. The OAF programs recommended positive amounts of lime to all farmers. KALRO recommended lime to farmers if their soil was acidic, but since the program took place in an acidic region, we assume purchasing lime is equivalent to following lime recommendations for this sample.

Figure 3 and table 3, row 3 show the meta-analytic results. The odds ratio for following the lime recommendation in the season in which the programs were implemented is 1.21 (95% CI 1.13 to 1.28).[14] We cannot reject the null of homogeneous treatment effects. Table 4 shows the result pooling data from all the experiments. With this approach, we find that the odds ratio

[14]This includes survey results for KALRO (we only have survey data for the concurrent season) and administrative data for all other programs

for following lime recommendations is 1.13 (column 1).

For completeness, appendix table F2 shows results for each experiment using both a linear probability model and a logistic model.

5.3 Purchase of Fertilizers

Next we examine the impact of these programs on the use of chemical fertilizers at topdressing. Topdressing fertilizer use was recommended by four programs (KALRO, IPA/PAD1-K, IPA/PAD2-K, OAF2-K).[15] Each program recommended different types of fertilizer. Based on data availability, we define the outcome "following fertilizer recommendations" as using one of the types of topdressing fertilizers recommended (as measured by KALRO's survey data), or purchasing the recommended topdressing fertilizer (as measured by the OAF's and IPA/PAD's administrative data).

Combining the results from these four experiments in a meta-analysis, we find that for the odds ratio increase in the likelihood of following the fertilizer recommendations is 1.13 (95% CI 0.85 to 1.51) (table 3, row 4). Figure 4 shows the corresponding forest plot. We reject the null of homogeneous effects for the first season (Q statistic p-value 0.03, I^2=67.8% (CI 6.5% to 89%)). Pooling data from all the experiments the odds ratio of following fertilizer recommendations is 1.09 and statistically significant table (4, panel B column 1). Appendix table F4 shows results for each experiment.

5.4 Combined Effect on All Recommended Inputs

In addition to lime and topdressing fertilizer, some of the program recommended other inputs including planting fertilizer, hybrid seeds, and others. Table D1 reports the list of the inputs recommended (and measured) for each of the program.

In order obtain an estimate of the overall effect of the program that takes into account all the possible outcomes, we follow Borenstein et al. (2009) and conduct a meta-analysis with multiple outcomes per program. For each program, we calculate the average effect size as the average of

[15] Messages sent for OAF1-K and OAF3-R only focused on lime, so we do not include them in this section.

the outcome specific log odds, and derive its standard errors by assuming 0.5 correlation across effect sizes.[16] Table 3 row 5 reports the results. The estimated odds ratio is 1.22 (95% CI 1.16 to 1.28). Figure 5 (panel a) reports the corresponding forest plot.

As an alternative strategy, we standardize treatment effects following the construction of indices as per Kling et al. (2007).[17] The overall effect of the programs, expressed in terms of standard deviations, is 0.04 (95% CI 0.03 to 0.06) (table 3 row 7 and figure 6, panel a). Appendix table F5 shows the results by experiment (column 1).

5.5 Effect on Other Inputs

We also test whether the program had any effect on inputs that were not explicitly recommended as part of SMS programs. As for the inputs recommended, we perform two meta-analyses. First, we aggregate effect sizes for each of the program as the average of the outcome specific effects. Second, we aggregate all inputs in indices following Kling et al. (2007). Both approaches lead to small and statistically insignificant results (table 3 row 6 and 8). Figures 5 and 6 (panel b) report the corresponding forest plots.

The list of the inputs considered is reported in appendix table D1. Appendix table F5 shows results (expressed in indices) for each program. In all cases the coefficients are small and statistically insignificant.

5.6 Treatment Persistence & Re-Treatment

Table F3 reports the results on following lime recommendations in the second season. In two of the programs (KALRO and IPA/PAD2the -K), farmers received only one round of messages, while in other programs messages were sent also prior to the following agricultural season. We find significant effects on the second season for the IPA/PAD programs. The results from OAF3-R also show positive results but not for OAF1-K nor OAF2-K.

[16]Our results are robust to different assumptions on the correlation across outcomes, including 0 and 1.

[17]We use the seemingly-unrelated regression framework to account to covariance across estimates.

5.7 Are there information spillovers?

Five out of the six projects rely on individual randomization to estimate impacts. This approach has more statistical power than group randomization, but it does not account for spillovers. If farmers who received messages shared the information with farmers in the control group, we would underestimate the impacts of these interventions. There is a growing literature suggesting that agricultural information diffusion is impeded by various frictions (Magruder, 2018) and previous qualitative and quantitative work in western Kenya suggests that these frictions might be important (Fabregas et al., 2017a; Duflo et al., 2008).

The randomization protocol for OAF3-R allows us to measure spillovers. In that experiment, a subset of groups (G3) were randomly assigned to be partially treated, with only half of the farmers receiving text messages, and a subset of groups was assigned to pure control (G0). This allows us to compare outcomes for farmers who did not receive messages but were in groups with farmers who did, against farmers in the control group. We find that that subset of farmers were nearly 0.4 percentage points (10%) more likely to purchase lime than those in the control group (table F7, panel C, column (2)). These point estimates are about half of the effect of direct treatment. These results suggest that the effects we measure are likely to be an underestimate of the effects of the programs. In addition, the presence of information spillovers also makes a much stronger case for public provision of these programs. A separate paper by Harigaya et al. (2018) provides additional information on the spillovers effects of the OAF3-R program.

Although the other programs were not designed to estimate spillovers across farmers, the variation in the number of treated farmers within OAF groups might also allow us to estimate spillovers. Table F7, columns (1) and (5) estimate whether having one additional treated farmer in the group increases the probability of purchasing lime. We find a small positive effect for OAF2-K but no effect for OAF1-K and OAF3-R.

Finally, for the OAF3-R program, we test for spillover effects on farmers that did not have a valid phone number (columns (3), (4), (7), and (8)). We find evidence of spillovers for this group.

5.8 Self-reported vs Administrative data

We collect both survey and administrative data in three out of the six projects (IPA/PAD1-K,IPA/PAD2-K,OAF1-K). Tables F2 shows some discrepancies between these two sources of data. For OAF1-k, the survey data lined up well with the administrative reports. However, for the IPA/PAD programs, particularly the second program, the survey results were statistically larger than the ones estimated using data from coupon redemption. This could indicate two things. One possibility is that the survey data is affected by social desirability or recall bias, and that true lime purchases are misreported. A second possibility is that the coupon redemption underestimates true lime use, since farmers might have acquired lime from other sources. We explore these possibilities for farmers in the IPA/PAD2-K sample. First, we check whether those farmers who were more likely to have other sources of lime (because they also reported participating in OAF programs) are more likely to report using lime but not redeeming the coupon. We find that within this sample, participating in OAF programs (35% of the sample) is associated with a 4 percentage point increase in the likelihood of reporting using lime in the survey but not redeeming the coupon (from 8 to 12%). This suggests that farmers can acquire lime from sources other than the shops that participated in the study. Second, we contrast the information from a survey completed with agricultural supply dealers in the region about the products they stock against farmer self-reported information about sources of lime they reported using. We find that only 36% of farmers who report using lime (but who did not redeem the coupon) mentioned that they had acquired lime from a shop that had also reported stocking lime during that period.[18] This could suggest that there is some degree of misreporting in the survey data. Overall, we believe that true effects are likely to be between these two bounds.

[18] If we impute a zero for those farmers who reported using lime, but who who did not redeem coupon and reported obtaining the lime from a shop that did not stock it, the coefficient from the linear probability model is 0.05 with s.e. of 0.011).

6 Discussion

Can SMS effects be strengthened by phone calls? Using one-way text messages to reach farmers may limit the amount of information that can be conveyed. For instance, one could hypothesize that text messages are effective at increasing awareness about the use of lime, but that farmers need more detailed information before adopting.[19] Speaking directly with extension personnel, would allow farmers to ask specific questions or clarify the content of the messages. We find low demand for this type of service. For instance, treated farmers in the OAF experiment were also given access to a toll-free number they could use to clarify the information, but less than 1% of the sample used it.

Low demand does not rule out the possibility that if farmers had received a call they would have changed their behavior. To explore this issue, we look at the differential effects of each treatment arm of the PAD/IPA2-K experiment. The experiment had three arms: in the first arm farmers only received text messages (SMS), in the second arm farmers received text messages and a phone call from an extension officer who explained the content of the text message (SMS+Call). In the third arm farmers received text messages and were offered the possibility to receive a call (SMS+Call Offer). We do not find statistically significant differences between any of the treatment arms (panel B in table G1 in the appendix), suggesting that receiving the call did not make a large difference in behavior.[20]

Who is most responsive to the programs? We find little systematic heterogeneous effects by gender, levels of education, farm size, age, and whether farmers used or heard about the inputs in the past (table 5). We show results for each program individually in table H1 and table H2. A particular important aspect of this lack of heterogeneity is that we do not find differential impacts for those who already knew about agricultural lime. This might suggest that extension programs might work through channels other than just increasing awareness about the technologies (Emerick et al., 2016).

[19]For instance, in other contexts, like political canvassing, interactions with real people have been more effective than other methods to mobilize people (Gerber and Green, 2000)

[20]If anything the point estimate for an offer to receive a call was higher then that of receiving the call. The p-value of this difference is 0.11.

Effects of Message Repetition. The advertisement literature suggests that the effects of advertisement repetition follows an inverted U-shaped curve (Park et al., 2008). Table 4 shows the effect of receiving one additional messages pooling data for all the programs and exploiting differences in the number of messages both across programs and within program. The results suggest that one additional message increases the odds ratio of following lime recommendations by 1.01 (panel A, column 2). In particular, one additional message mentioning lime increases the odds ratio of 1.03, while an additional message mentioning fertilizer has no effect and on the odds ratio of following lime recommendations (column 3). We find similar results for top-dressing fertilizer (panel B). In the OAF2-K and OAF3-R experiments the number of messages was randomized across farmers, which provides us with the opportunity to look at this question in more details. We find that receiving one additional SMS increases the odds ratio of purchasing lime by 1.03 in the OAF2-K program and 1.07 in the OAF3-R program (table G2, column 4). In both programs the effect is driven by receiving at least 2 SMS messages and we find no effect from additional messages (columns 5 and 6).

Effects of Message Content and Framing. Whether using behavioral insights to frame messages or providing additional information to farmers can strengthen impacts is an important policy question. We use the experimental design of four experiments to draw some lessons on framing and content. Table G3 in the appendix shows effects for different types of messages experimentally sent in the OAF2-K and OAF3-R projects. OAF2-K randomized five types of messages in addition to the general recommendation. These included highlighting potential for yield increases, encouraging own experimentation, encouraging experimentation with neighbors, messages that highlighted social comparisons, and self-efficacy. The OAF3-R experiment randomized six different types of messages (yield impacts, self diagnosis, focus on soil test, focus on how lime works, nudge to order immediately, mention soil acidity). We cannot reject the hypothesis that all messages were equally effective at changing the probability of ordering lime. The only statistically significant effect we detect is for the OAF2-K program, where we find that messages targeting the family instead of the individual farmer reduced the probability of purchasing lime by 1.6 percentage points.

Using the experimental variation from different treatment arms in IPA/PAD1-K, we can estimate whether the treatment arms that provided additional information about the extent of soil acidity in the area were more effective. We find that the point estimates are larger for the treatment arm that provided more information about acidity but they are not statistically different from each other (panel A in table G1). Similarly, one of the treatment arms from the OAF1-K program also provided specific information about soil acidity, but we do not find statistically significant differences across treatment arms (panel C). Overall, these results suggest that the way in which the messages were framed made little difference. However, since power is limited and the cost of optimizing messages is very low, this is an area that warrants further exploration.

7 Cost Analysis

Ideally, one would want to estimate the rate of return of these programs to judge whether these programs are worth the investment. Producing reliable estimates on the returns to SMS interventions is difficult since the effects are modest, the cost of delivery is very low, and outcome variables like crop yields and profits tend to be extremely noisy.[21] Our experiments were not designed nor powered to detect impacts on yields.

To give a sense of the returns to these programs, we present two types of calculations. First, since the main purpose of the programs was to use SMS to encourage experimentation with agricultural lime, we calculate the average cost per farmer following lime recommendations due to a program of this type, and compare it with the cost of other programs pursuing the same objective with in-person approaches. This is not sufficient to inform the overall investment decision, but if we take the policy objective as given, it is useful as a point of comparison with other extension approaches. Second, we conduct back-of-the-envelope calculations to provide some estimates of the benefit-cost ratio of an intervention of this kind implemented at scale. To establish benefits, we combine information from the point estimates from the increased quantity

[21] This problem is similar to that of measuring the rates of returns to online advertising. Intermediate metrics such as clicks are common in measure effectiveness (Lewis and Rao, 2015).

of lime adopted with existing agronomic data that allows us to estimate the corresponding effect on yields. For the cost estimations we only consider the social marginal costs of the text messages.[22]

Cost-Effectiveness. The average impact of the SMS programs on the probability of following lime recommendation was 2 percentage points (table 3, LPM result). Since the programs involved, on average, four SMS messages, three of which specifically about lime, we estimate that the cost of getting one farmer to experiment with lime ranged between $0.15 and $0.20 USD if implemented at scale, assuming a unit cost of $0.001 per SMS.

We can compare these estimates to that of in-person extension approaches implemented by KALRO and OAF. In particular, we have experimental estimates that KALRO's Farmer Field Days (FFDs), large meetings with farmers, increased use of agricultural lime by 3.6 percentage points (Fabregas et al., 2017a). Based on information reported by KALRO, we calculate that each FFD cost about $2,600 to implement. This includes all costs for staff, transport, compensation and materials required to set up the test plots, invite presenters, advertise the FFDs to farmers and carry out the events. Since each FFD hosted between 100 and 300 farmers, this amounts to a per farmer cost of at least $9.[23]. Given that the FFDs covered various topics we attribute 1/5 of their cost to the lime program and estimate that the cost per adopting farmer was $50.

Between 2016 and 2017, One Acre Fund conducted a randomized control trial in western Kenya with the objective to test alternative ways of encouraging experimentation with lime. They tested a combination of sales incentive for field officers, the extension agents in charge of signing contracts with farmers and providing training, and free lime for farmers and field officers. OAF found that the sales incentives alone increased lime adoption by 13.4 percentage points and were by far the most cost-effective approach among those tested as part of the trial. This program involved a payment of $0.5 per adopting farmer, up to a maximum of $60 per

[22]This assumes that other potential fixed costs of running these programs would be incurred with or without the SMS component, and the social costs of text messages rather than the costs incurred by an organization purchasing these services from a telecommunications company.

[23]In India, (Emerick et al., 2016) estimates a per farmer cost of field days are about $5 US ($200 per FFD attended by 41 farmers) However, they estimate that their FFD generated one-year revenue gains of $410.

field officer, plus a day of training for the field officers. We estimate that the cost per adopting farmer was $1.89.[24]

Cost-Benefit. The median effect of four agronomic trials measuring the impact of lime application on yields, was 1.4 kg per kg of lime applied.[25] The cost of applying 1 additional kg of lime is estimated to be approximately $0.15, which takes into account the local price of lime ($0.10/kg) and transport and application cost ($0.05/kg).[26] The revenue obtained from one additional kg of maize is assumed $0.35, which takes into account the market price of maize ($0.40/kg), minus additional costs for harvesting and transport ($0.05/kg).[27] Therefore, the profits from applying one kg of lime is estimated to be approximately $0.33.

The overall impact of the programs in terms of quantity of lime applied was found to be 1.24 kg (table 3), which implies a benefit of $0.41 per farmer treated. Considering that the cost of the programs was on average $0.04 per farmer, the benefit-cost ratio is 10. However, at scale, with a unit cost of $0.001 per SMS, the implied benefit-cost ratio would be 100.

8 Conclusion

The spread of cellphones in developing countries has opened new opportunities to reach farmers with timely and customized agricultural extension information.

We evaluate the effects of six different programs implemented in Kenya and Rwanda, finding that they cost-effectively changed farmer behavior. We do not find evidence to suggest that providing additional details about local soil characteristics or changing the framing of the messages made a difference. There was low demand for additional services, such as access to a call center, and we have some evidence that indicates that additional calls did not make a difference

[24]We assume that the cost of the training for the 23 field officers participating in the program was $1,000. The cost of the sales incentives was calculated assuming that 15.4% of the 5727 farmers treated purchased lime (the control group mean was 2%, and that the maximum incentive of $60 was never reached.

[25]This is the median effect of four studies aimed at estimating the effect of micro-dosing lime on maize yield. The impact per kg of lime applied was: 1.8 kg (OAF, 2014), 2.47 kg (OAF, 2015), 0.23 kg (Kisinyo et al., 2015), and 0.99 kg (Omenyo et al., 2018)

[26]The price of lime reflects the average price of lime in western Kenya during the 2017 main agricultural season, based on data collected by IPA-K.

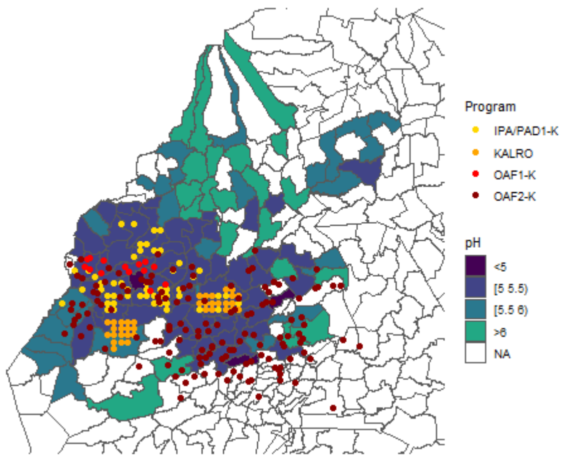
[27]Maize prices and assumed costs are based on data collected by IPA-K in the study area during the 2017 main agricultural season.

in adoption. However, we find that repeating the same message had a statistically significant impact on the adoption of inputs.

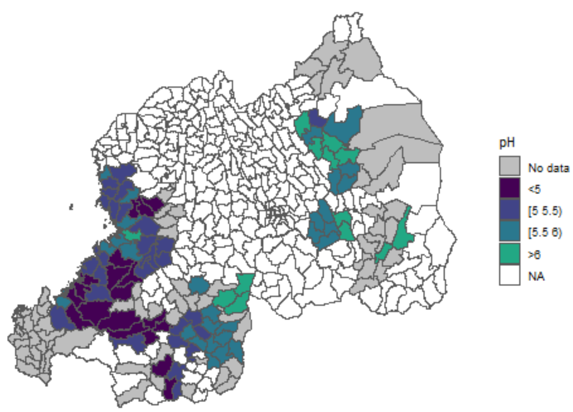
Our back-of-the envelope calculations suggest that text-based approaches can be cost-effective from the point of view of a principal who is interested in promoting new inputs.

9 Figures and Tables

Figure 1: Project Maps



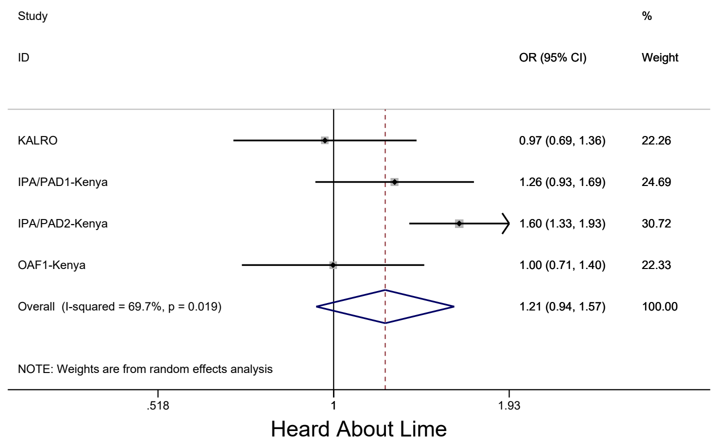
(a) Western Kenya



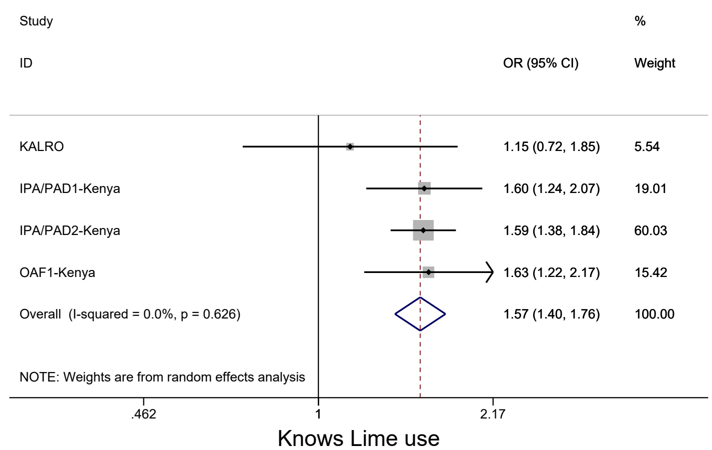
(b) Rwanda

Notes: Panel (a) shows the median level of pH in all wards in which the IPA/PAD2-K program took place as well and the location of the other programs. Panel (b) shows the sectors in which the OAF3-R program took place and the median level of pH, where available.

Figure 2: Combined Effects on Knowledge About Lime



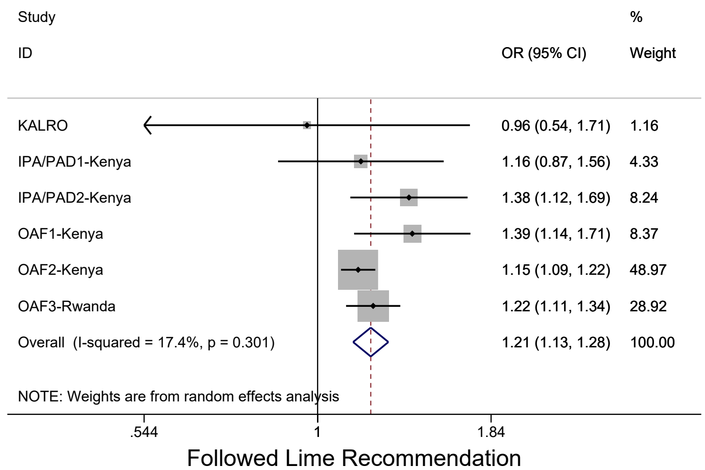
(a) "Have you heard about lime?"



(b) Mentions lime as a way to reduce acidity

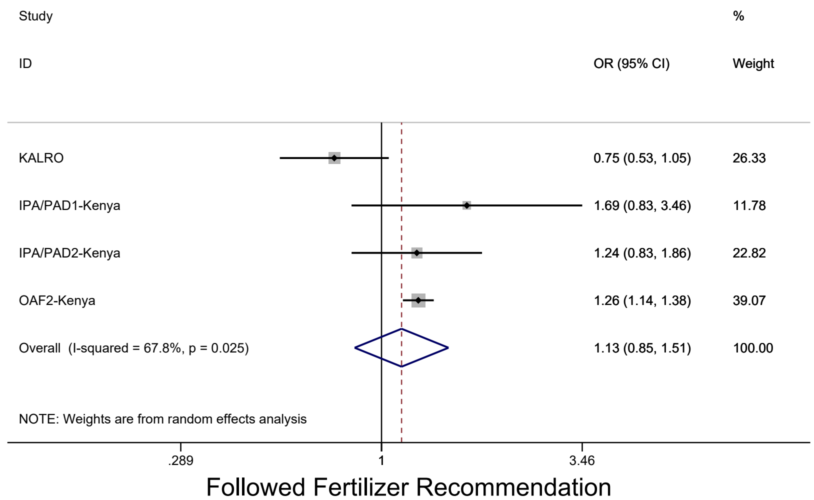
Notes: The figure plots the meta-analysis results for specific outcomes. The effects are estimated using a random-effects meta-analysis model. Results are reported in odds ratios. The horizontal lines denote 95% confidence intervals.

Figure 3: Combined Effects on Lime



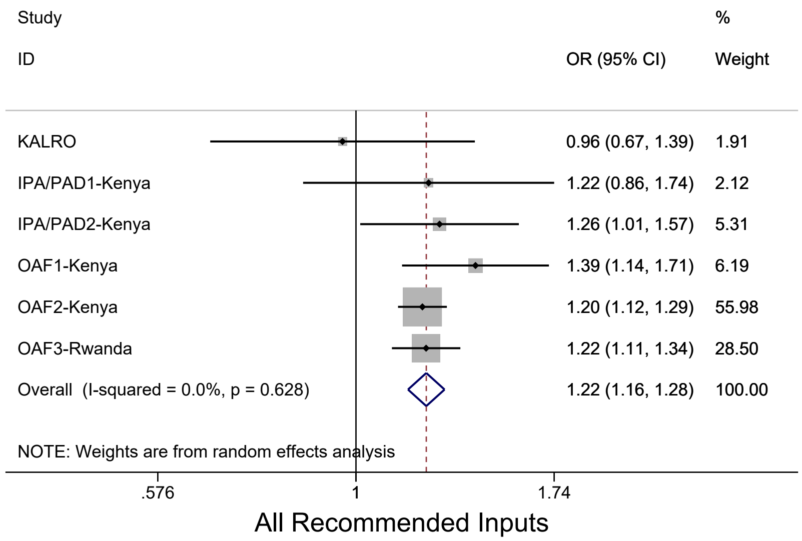
Notes: The figure plots the meta-analysis results for following lime recommendations. The effects are estimated using a random-effects meta-analysis model. Results are reported in odds ratios. The horizontal lines denote 95% confidence intervals. The KALRO results are measured using survey data, while all others are measured using administrative data.

Figure 4: Combined Effects on Fertilizer

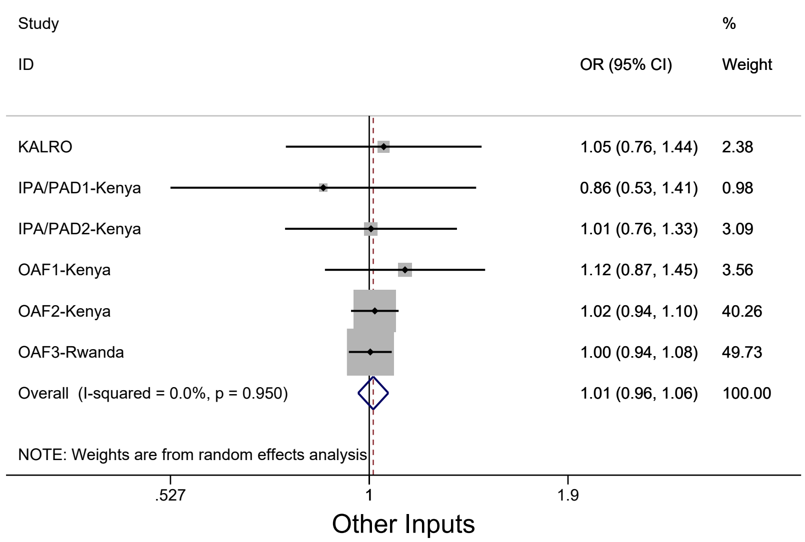


Notes: The figure plots the meta-analysis results for following fertilizer recommendations. The effects are estimated using a random-effects meta-analysis model. Results are reported in odds ratios. The horizontal lines denote 95% confidence intervals. The KALRO results are measured using survey data, while all others are measured using administrative data.

Figure 5: Combined Effects on Recommended and Other Inputs (Odds Ratios)



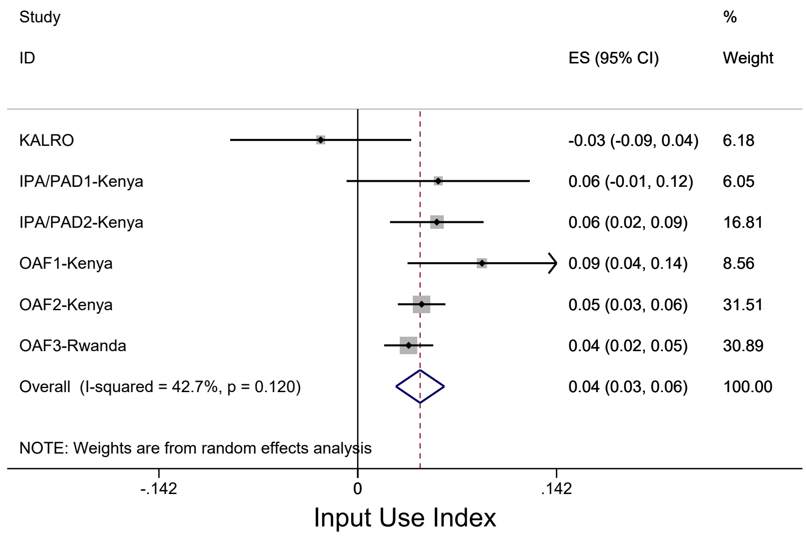
(a) Recommended Inputs



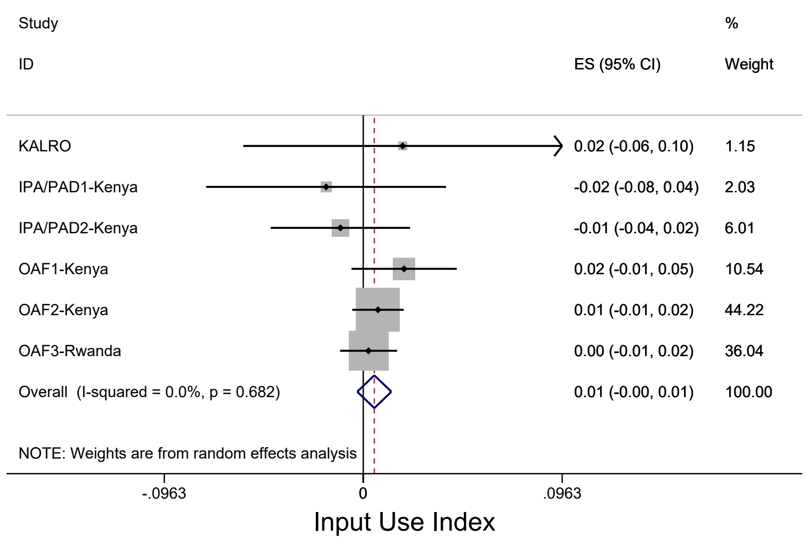
(b) Other Inputs

Notes: The figure plots the meta-analysis results for the effect of the programs on use or purchases of recommended inputs and other inputs not mentioned by the SMS messages. The effects are estimated using a random-effects meta-analysis model. Multiple outcomes per study are aggregated assuming that correlation across outcomes is equal 0.5. Results are reported in standard deviations. The horizontal lines denote 95% confidence intervals. Panel (a) reports results for recommended inputs. Panel (b) reports results for other inputs.

Figure 6: Combined Effects on Recommended and Other Inputs (Index)



(a) Recommended Inputs



(b) Other Inputs

Notes: The figure plots the meta-analysis results for the effect of the programs on use or purchases of recommended inputs and other inputs not mentioned by the SMS messages. The effects are estimated using a random-effects meta-analysis model. Results are reported in standard deviations. The horizontal lines denote 95% confidence intervals. Panel (a) reports results for recommended inputs. Panel (b) reports results for other inputs.

Table 3: Meta-analysis Results (Random Effects)

		Effects				Heterogeneity				
Row #	Outcome	N	Effect	95% CI		Q stat (p-value)	I^2	I^2 - 95% CI		τ^2
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Odds Ratios									
1	Heard Lime	4	1.21	0.94	1.57	0.02	69.69	12.76	89.47	0.05
2	Knowledge Acidity	4	1.57	1.40	1.76	0.63	0.00	0.00	84.69	0.00
3	Lime Rec.	6	1.21	1.13	1.28	0.30	17.45	0.00	62.24	0.00
4	Fertilizer Rec.	4	1.13	0.85	1.51	0.03	67.84	6.48	88.94	0.05
5	All Recommended Inputs	6	1.22	1.16	1.28	0.63	0.00	0.00	74.62	0.00
6	Other Inputs	6	1.01	0.96	1.06	0.95	0.00	0.00	74.62	0.00
	Standard Deviations									
7	Recomm Inputs (Index)	6	0.04	0.03	0.06	0.12	42.73	0.00	77.33	0.00
8	Other Inputs (Index)	6	0.01	-0.00	0.01	0.68	0.00	0.00	74.62	0.00
	Kg									
9	Kg Lime	5	1.26	0.26	2.27	0.00	85.96	69.23	93.60	1.06
	LPM									
10	Lime Rec.	6	0.02	0.01	0.03	0.00	72.18	35.75	87.96	0.00

Notes: Meta-analysis results for each outcome reported in the rows. Column (2) -(5) reports results from a random-effects model; Column (6)-(9) reports heterogeneity results. The coefficient represents the estimated summarized effects across studies, measured in odds ratios (except for the 'Recomm Inputs' and the 'Other Inputs' variable, which are index variables and are measured in standard deviations, 'Kg Lime', which is measured in kg, and the LPM results for following lime recommendations).

Table 4: Pooled Regressions

	Odds ratios		
	(1)	(2)	(3)
Panel A. Followed Lime Recommendations			
Treated	1.137***		
	(0.026)		
N SMS (tot)		1.014***	
		(0.002)	
N lime SMS			1.031***
			(0.005)
N fert SMS			1.000
			(0.008)
Mean Control	0.14	0.14	0.14
Observations	132065	132065	132065
Panel B. Followed Fertilizer Recommendations			
Treated	1.095**		
	(0.041)		
N SMS (tot)		1.011**	
		(0.004)	
N lime SMS			1.016*
			(0.009)
N fert SMS			1.026**
			(0.011)
Mean Control	0.13	0.13	0.13
Observations	41132	41132	41132

Notes: This table the effect of the programs on following lime (panel A) and fertilizer recommendations (panel B) analysis pooling data from all programs. Both dependent variables are measured using administrative data for all programs except for KALRO, where survey data is used. All regressions include program FEs. Effect sizes are reported in terms of odds ratios measured using Logit. Bootstrap standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 5: Heterogeneity (Pooled)

	Odds ratios					
	Female	Primary School	Large Farm	Young	Used Input	Heard Input
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Followed Lime Recommendations						
Treated	1.176***	1.119	1.130***	1.113***	1.144***	1.135**
	(0.048)	(0.101)	(0.028)	(0.033)	(0.042)	(0.072)
[X]	1.316***	1.068	1.341***	0.705***	3.264***	1.069
	(0.060)	(0.098)	(0.057)	(0.032)	(0.259)	(0.133)
[X] *Treated	0.948	1.058	1.021	1.039	1.002	1.021
	(0.049)	(0.122)	(0.052)	(0.057)	(0.101)	(0.152)
Mean Control	0.29	0.23	0.14	0.31	0.07	0.25
Observations	44969	9711	132065	41315	94609	8560
Panel B. Followed Fertilizer Recommendations)						
Treated	1.034	0.947	1.098**	1.095*	1.088**	1.217
	(0.068)	(0.168)	(0.049)	(0.059)	(0.045)	(2.395)
[X]	1.135**	1.307*	1.077	0.777***	8.976***	11.352
	(0.071)	(0.211)	(0.073)	(0.041)	(0.530)	(22.921)
[X] *Treated	1.085	1.078	0.989	1.005	1.078	0.654
	(0.074)	(0.221)	(0.078)	(0.067)	(0.064)	(1.292)
Mean Control	0.13	0.10	0.13	0.13	0.13	0.50
Observations	40157	8560	41132	40164	41132	773

Notes: This table shows results of heterogeneity analysis pooling data from all programs. The dependent variable is whether the farmer followed lime recommendations (panel A) or fertilizer recommendations (panel B) in the first season. Both dependent variables are measured using administrative data for all programs except for KALRO, where survey data is used. We show results for gender, whether respondent completed primary school, whether the respondent's land is large (defined as above median use of inputs for the OAF samples and more than 1.5 acres of land for the other programs), whether the respondent was under 40 years old, whether the respondent had previously used the input, and whether the respondent had previous knowledge of the input. All regressions include program FEs. Effect sizes are reported in terms of odds ratios measured using Logit. Bootstrap standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

References

Aker, J. C. (2017). Using digital technology for public service provision in developing countries. *Digital Revolutions in Public Finance*, pages 201–24.

Aker, J. C., Collier, P., and Vicente, P. C. (2017). Is information power? using mobile phones and free newspapers during an election in Mozambique. *Review of Economics and Statistics*, 99(2):185–200.

Aker, J. C., Ghosh, I., and Burrell, J. (2016). The promise (and pitfalls) of ICT for agriculture initiatives. *Agricultural Economics*, 47(S1):35–48.

Aker, J. C., Ksoll, C., and Lybbert, T. J. (2012). Can mobile phones improve learning? evidence from a field experiment in Niger. *American Economic Journal: Applied Economics*, 4(4):94–120.

Aker, J. C. and Mbiti, I. M. (2010). Mobile phones and economic development in Africa. *Journal of Economic Perspectives*, 24(3):207–32.

Anderson, J. and Feder, G. (2007). Agricultural extension. *Handbook of agricultural economics.*

Borenstein, M., Hedges, L., Higgins, J., and HR, R. (2009). *Introduction to Meta-Analysis*. John Wiley.

Borenstein, M., Higgins, J. P., Hedges, L. V., and Rothstein, H. R. (2017). Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Research synthesis methods*, 8(1):5–18.

Brady, N. and Weil, R. (2004). Soil acidity, alkalinity and salinity. *Elements of the Nature and Properties of Soils*, pages 266–315.

Buntaine, M. T., Jablonski, R., Nielson, D. L., and Pickering, P. M. (2018). Sms texts on corruption help ugandan voters hold elected councillors accountable at the polls. *Proceedings of the National Academy of Sciences*, 115(26):6668–6673.

Camacho, A. and Conover, E. (2010). The impact of receiving price and climate information in the agricultural sector. *Working Paper.*

Casaburi, L., Kremer, M., Mullainathan, S., and Ramrattan, R. (2014). Harnessing ICT to increase agricultural production: Evidence from Kenya. *Working Paper.*

Cole, S. A. and Fernando, A. N. (2016). 'mobile'izing agricultural advice: Technology adoption, diffusion and sustainability. *Working Paper*.

Cunha, N., Lichand, G., Madeira, R., and Bettinger, E. (2017). What is it about communicating with parents? *Working Paper*.

DerSimonian, R. and Laird, N. (1986). Meta-analysis in clinical trials. *Controlled clinical trials*, 7(3):177–188.

Duflo, E., Kremer, M., and Robinson, J. (2008). How high are rates of return to fertilizer? evidence from field experiments in Kenya. *The American Economic Review*, 98(2):482–488.

Duflo, E., Kremer, M., Robinson, J., and Schilbach, F. (2018). Learning about appropriate technology use: Evidence from fertilizer use in western Kenya. *Working Paper*.

Dustan, A., Maldonado, S., and Hernandez-Agramonte, J. M. (2018). Motivating bureaucrats with non-monetary incentives when state capacity is weak: Evidence from large-scale field experiments in peru. *Working Paper*.

Emerick, K., de Janvry, A., Sadoulet, E., and Dar, M. H. (2016). Identifying early adopters, enhancing learning, and the diffusion of agricultural technology.

Evenson, R. E. and Gollin, D. (2003). Assessing the impact of the green revolution, 1960 to 2000. *Science*, 300(5620):758–762.

Fabregas, R., Kremer, M., Robinson, J., and Schilbach, F. (2017a). The effectiveness of public agricultural extension: Evidence from two approaches in Kenya. *Working Paper*.

Fabregas, R., Kremer, M., Robinson, J., and Schilbach, F. (2017b). Netflix for agriculture. *Working Paper*.

Fafchamps, M. and Minten, B. (2012). Impact of sms-based agricultural information on Indian farmers. *The World Bank Economic Review*, 26(3):383–414.

FURP (1995). New fertilizer recommendations for small-scale farmers in kenya. *KARI Information Bulletin - No. 12*.

George, S., Gupta, S., and Neggers, Y. (2018). Coordinating voters against criminal politicians: Evidence from a mobile experiment in india.

Gerber, A. S. and Green, D. P. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American political science review*, 94(3):653–663.

Gudu, S., Okalebo, J., Othieno, C., Obura, P., Ligeyo, D., Shulze, D., and Johston, C. (2005). Response of five maize genotypes to nitrogen, phosphorus and lime on acid soils of Western Kenya. In *African Crop Science Conference Proceedings*, volume 7, pages 1109–1115.

Harigaya, T., Kremer, M., Lowes, M., On, R., and Zane, G. (2018). Behavioral spillovers via sms: Experimental evidence from Rwanda. *Working Paper*.

Higgins, J. P., Green, S., et al. (2008). Cochrane handbook for systematic reviews of interventions.

Higgins, J. P. and Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in medicine*, 21(11):1539–1558.

Higgins, J. P., Thompson, S. G., Deeks, J. J., and Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj*, 327(7414):557–560.

IPNI (1999). Nitrogen notes. *Manual, http://www.ipni.net/NitrogenNotes*.

Karlan, D., McConnell, M., Mullainathan, S., and Zinman, J. (2016). Getting to the top of mind: How reminders increase saving. *Management Science*, 62(12):3393–3411.

Karlan, D., Morten, M., and Zinman, J. (2012). A personal touch: Text messaging for loan repayment. Technical report, National Bureau of Economic Research.

Kiplagat, J. K., Okalebo, J. R., Serrem, C. K., Mbakaya, D. S., and Jama, B. (2014). *Determination of Appropriate Rate and Mode of Lime Application on Acid Soils of Western Kenya: Targeting Small Scale Farmers*, pages 177–186. Springer International Publishing, Cham.

Kisinyo, P., Opala, P., Palapala, V., Gudu, S., Othieno, C., and Ouma, E. (2015). Micro-dosing of lime, phosphorus and nitrogen fertilizers effect on maize performance on an acid soil in Kenya. *Sustainable Agriculture Research*, 4(2):21.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Ksoll, C., Aker, J., Miller, D., Perez-Mendoza, K., and Smalley, S. (2014). Learning without teachers? a randomized experiment of a mobile phone-based adult education program in Los Angeles. *Working Paper*.

Larochelle, C., Alwang, J., Travis, E., Barrera, V. H., and Andrade, J. M. D. (2019). Did you really get the message? Using text reminders to stimulate adoption of agricultural technologies. *The Journal of Development Studies*, 55(4):548–564.

Lewis, R. A. and Rao, J. M. (2015). The unfavorable economics of measuring the returns to advertising. *The Quarterly Journal of Economics*, 130(4):1941–1973.

Magruder, J. R. (2018). An assessment of experimental evidence on agricultural technology adoption in developing countries. *Annual Review of Resource Economics*, 10:299–316.

Mortvedt, J. L. M. and Follet, R. (1999). Fertilizer technology and application. *Meister Publishing Co. Willoughby, OH*.

NAAIAP (2014). Soil suitability evaluation for maize production in Kenya. *Manual*.

Nakasone, E., Torero, M., and Minten, B. (2014). The power of information: The ICT revolution in agricultural development. *Annu. Rev. Resour. Econ.*, 6(1):533–550.

OAF (2014). Managing soil acidity with lime (2014 phase 2 trial report). Technical report, One Acre Fund.

OAF (2015). Managing soil acidity with lime (2015 trial report). Technical report, One Acre Fund.

OAF (2017). Annual report.

Omenyo, V. S., Okalebo, J., and Othieno, C. (2018). Effects of lime and phosphorus fertilizer on maize performance in acid soils of western Kenya. *Second RUFORUM Biennial Meeting 20 - 24 September 2010, Entebbe, Uganda*.

Overdahl, J. C., George, W. R., and Harvey, L. M. (2017). Fertilizer urea: Nutrient management. *University of Minnesota Extension*.

Park, T., Shenoy, R., and Salvendy, G. (2008). Effective advertising on mobile phones: a literature review and presentation of results from 53 case studies. *Behaviour & Information Technology*, 27(5):355–373.

Pauw, E. F. D. (1994). The management of acid soils in Africa. *Outlook on Agriculture*, 23(1):11–16.

Plaster, J. (2003). Soil science and management. *Delmar Learning, a division of Thomas Learning, Inc*, pages 319–322.

Terman, G. and Engelstad, O. (1976). Agronomic evaluation of fertilizers: Principles and practices. *Muscle Shoals, Ala.: Tennessee Valley Authority*.

The, C., Calba, H., Zonkeng, C.and Ngonkeu, E. L. M., and Adetimirin, V. O. (2006). Response of maize grain yield to changes in acid soil characteristics after soil amendment. *Plant Soil*.

Tisdale, S. L., Nelson, W. L., and Beaton, J. D. (1990). Soil fertility and fertilizers. *New York: Macmillan*.

Tjernstrom, E., M., C., and Lybbert, T. J. (2015). The dirt on dirt: Soil characteristics and variable fertilizer returns in Kenyan maize systems. working paper. Technical report, Working Paper.

Van Campenhout, B., Spielman, D. J., and Lecoutere, E. (2018). Information and communication technologies (icts) to provide agricultural advice to smallholder farmers experimental evidence from Uganda. *IFPRI Discussion Paper*.

Wasonga, C. J., Sigunga, D. O., and Musundu, A. O. (2008). Phosphorus requirements by maize varieties in different soil types of Western Kenya. *Africa Crop Science Journal*, vol. 16:161–173.

World Bank (2012). Information and communications for development 2012: Maximizing mobile. Technical report.

Appendix

Table of Contents

A Description of Programs and Experiments

A.1 KALRO's Program

The Kenya Agriculture and Livestock Research Organization (KALRO) is a public agency with the mandate to promote agricultural research and dissemination in Kenya. In 2014 and 2015, KALRO's Kakamega office implemented two extension programs aimed at encouraging smallholder farmers to adopt inputs and management practices that could address some of the regional soil deficiencies. This programs reflected their goal of reaching a large number of farmers at a lower cost than that of in-person individual farm visits.[28]

KALRO's SMS program consisted of sending 20 different agriculture-related text messages to maize farmers' mobile phones. The content of the messages was developed by the Ministry of Agriculture, Livestock and Fisheries and the delivery was managed by KALRO.[29] Each message provided broad advice on best practices but most messages did not provide actionable advice on agricultural practices. For instance they encouraged farmers to "*buy recommended certified maize and legume seed from approved agrodealers*" and "*obtain information on favorable market prices before you sell your harvest*". One message advised farmers to test their soil's pH, and another one recommended farmers to use lime if their soil is acidic, stating: "*if the soil is acidic (pH less than 5.5), apply recommended rate of agricultural lime at least 30 days before planting*". The message also provided farmers with phone numbers where they could inquire about purchasing a soil test to assess their farms' pH. Appendix B provides additional details and lists all the messages sent during the intervention.

KALRO's program was evaluated in partnership with IPA. To recruit farmers into the pro-

[28]KALRO experimented with two approaches. First, with farmer field days (FFD), one-day events in which a large number of farmers can observe demonstration plots and receive information from extension agents. The second approach consisted of delivering agricultural messages to farmers via SMS. This paper focuses on the results from the second approach, but we discuss further details of the impacts of FFDs in Fabregas et al. (2017a) and in appendix B.

[29]Since 2014 the Ministry of Agriculture has announced plans to roll out an e-extension system to reach over 7 million farmers, by providing phone-based support to extension workers who would then advise farmers. The version of the program that was evaluated was a pilot program that tried to deliver information directly to farmers. In July 2018, the Kenyan Ministry of Agriculture and Irrigation, in partnership with PAD and Safaricom, launched an SMS service (MoA-Info) aimed at providing agricultural advice to farmers across the country.

gram the evaluation team conducted a census of farmers in the Ugenya and Mumias sub-counties using specific walking rules to visit a representative sample of households. Farmers who owned a mobile phone, had grown maize or legumes during the previous year, and were in charge of farming activities in the household were then invited to participate in the project.[30]

In September 2014, farmers completed an in-person baseline survey and were then randomized into the SMS treatment (415 farmers) or a comparison group (417 farmers).[31] Table E1 reports summary statistics, showing balance for observed characteristics. The text-message service was implemented between July 2015 and November 2015, in the period corresponding to the short rains season. An in-person endline survey, asking information about input use and knowledge, was completed with 92% of the baseline sample by January 2016. We do not find evidence of differential attrition by treatment group (appendix table E7, panel A).

At the end of the endline survey, all farmers received two (paper) discount coupons redeemable at selected agricultural supply dealers in their nearest market center. The coupons were devised as a way to collect information on input choices and reduce concerns about enumerator demand effects since purchasing decisions would be made at a later time when farmers were not directly observed by any member of the research or KALRO teams. The first discount coupon was redeemable for a 50% discount for agricultural lime. The second coupon was redeemable for a 50% discount for any chemical fertilizer of their choice (NPK, DAP, CAN, urea or Mavuno).[32] Coupon redemption was possible up to the start of the subsequent 2016 long rain agricultural season (March 2016). Participating agricultural supply dealers were instructed (and incentivized through a small payment) to keep clear records on input choices and quantities purchased by farmers who redeemed coupons.[33] Therefore, the questions of the endline survey measure behavior that occurred during the season when the program was implemented, whereas the coupons measure purchasing behavior that occurred the following agricultural season.

[30]Enumerators completed a total of 1,330 census surveys and approximately 94% of those recruited during census activities met the selection criteria.

[31]A third group of 417 farmers was randomized into the FFD program.

[32]Both coupons had an upper limit discount of approximately $10 USD.

[33]Incentives were paid on the basis of having both the physical coupon and a record of the purchase in their logbooks.

A.2 IPA & PAD's Programs

PAD is a non-profit organization that supports the provision of phone-based customized agricultural information services to smallholder farmers in developing countries. PAD supported two agricultural extension research projects in western Kenya that were implemented and evaluated by IPA. Both programs aimed to test different approaches to providing agricultural advice based on local soil information and encourage experimentation with agricultural lime and fertilizer.

A.2.1 Program 1 (IPA/PAD1-K)

Throughout the 2016 short rain agricultural season, IPA, with support from PAD, sent selected farmers text messages with information about agricultural inputs (including lime and chemical fertilizers) as well as other general agronomic recommendations on maize farming. Farmers who participated in this program were recruited through administrative farmer records of a large agribusiness in the region and from records of individuals who had participated in IPA's activities previously.[34] In July 2016, a random sample of farmers from both databases were contacted over the phone to invite them to participate in the study and complete a short phone-based baseline survey to determine eligibility. Farmers who were planning to plant maize in the 2016 short rains season, had a farm located within the intervention area, and expressed interested in receiving agricultural information over their phone were invited to participate.[35]

Two types of messages were tested: messages with general advice for the program area but that did not refer to local soil data (e.g. *"Lime reduces soil acidity and makes nutrients such as phosphorous available to your maize"*) and messages that provided information from local soil tests (e.g. *"Based on soil tests performed around [area] we recommend you: apply [quantity] bottletop of lime and cover with soil and then apply [quantity] of DAP"*). Among farmers receiving these messages with specific information, those who lived in areas that had

[34] The Mumias Sugar Company ran a contract farming model with sugar cane farmers in the region up to 2015. The vast majority of farmers plants maize in addition to many sugar cane so the company supported delivery of maize extension messages. The farmers who appeared in the IPA database were mainly recruited through large school meetings, as discussed in Duflo et al. (2018). This group accounted for about 47% of the final sample.

[35] From 2,255 interviewed respondents, 2,131 consented to participate in the baseline. From that set 1,897 (89%) met the criteria for selection.

median pH of more than 5.5 did not receive message about lime (18% of the sample). Both groups of farmers also received messages about planting (DAP) and at topdressing (urea) fertilizers. Farmers received between 24 and 28 messages. Appendix B provides additional details and lists all the messages sent.

A final sample of 1,897 farmers was randomized into three groups: receiving the general messages, receiving the specific messages, and a control group. In addition, during the following agricultural season (long rains 2017) both treatment groups received five additional messages promoting the use of agricultural lime (both treatment groups received messages based on local soil characteristics). We show summary statistics and a full list of balance checks in table E2. We do not find evidence of systematic statistically significant differences between control and treatment groups at baseline.

We can measure impacts in two ways. First, two electronic discount coupons were sent via SMS to all participating farmers at the beginning of the season after the initial set of recommendations were sent.[36] All farmers, including those in the control group, received these coupons. The first coupon gave farmers a choice of either 10 kg of lime or 1 bar of soap. By allowing farmers to choose between lime and another common product of the same value, we intended to capture farmers' input choices without liquidity constraints. The second coupon, sent mid-season, provided a 30% discount on one type of top-dressing fertilizer (urea, CAN, or Mavuno), up to a pre-discount amount of 500 Ksh (approximately $ 5 USD). To redeem coupons, each farmer was assigned to an agricultural supply dealer in their preferred or closest market center (selected by farmers during baseline). To measure effects over a second season all farmers received a second round of lime coupons for the 2017 long rain season. This coupon only provided a 15% discount on the first seven 10-kg bags of agricultural lime. All farmers received a phone call around the time the coupon was sent, to ensure that treatment and controls were equally aware of the electronic coupon.

Second, we conducted a phone endline survey mid-2017 long rain season with the full sample

[36] The first coupon was sent 10 days after the beginning of the experiment, after 7 recommendation messages, with a reminder 1 week later. The second coupon was sent 1 month after the beginning of the experiment, after 18 messages, with a reminder after 10 days and another after 20 days.

of farmers participating in the experiment. The survey included questions about input use during the 2016 and 2017 agricultural seasons and farmers' general agricultural knowledge. Enumerators were able to survey 80% of farmers in the sample, and we do not find evidence of differential survey completion by treatment group (appendix table E7, panel B).

A.2.2 Program 2 (IPA/PAD2-K)

A second program was implemented the following agricultural season that incorporated lessons from IPA/PAD1-K and aim to test a low-cost way to recruit farmers through agricultural supply dealers. This recruitment method offered several advantages. First, it was a low-cost and quick method to recruit farmers: in a period of two months, over 8,000 were enrolled. Second, farmers who are clients of agricultural supply dealers might already be more likely to acquire inputs be less credit constrained and, therefore, benefit from an information-based program.

As part of this program, a total of 102 agricultural supply dealers in 46 market centers in Western Kenya invited farmers to enrol in a maize farmer census. The registration period ran from early December 2016 to late January 2017. All registered farmers were then contacted over the phone by a member of the research team to obtain consent to participate in the study and baseline information about their farming practices and previous input use. A total of 5,890 farmers completed the phone baseline survey, met the eligibility criteria, and resided in eligible areas for which PAD had soil information.[37]

Farmers were then randomized into four groups. The first three groups received PAD's SMS agricultural information services and the fourth group remained as a control. One third of treated farmers received information via SMS only, another third received SMS and were invited to express interest in receiving a phone call that would explain the messages, the last third of treated farmers were contacted over the phone and offered an explanation of the messages.

[37] A total of 8,496 farmers were registered through 144 agricultural supply dealers in 60 market centers. However, for logistical reasons the study area was later restricted to 46 market centers and 102 agricultural supply dealers. From that sample, farmers who were reached but did not complete the baseline survey included 257 who did not consent to participate in the study, 53 who were not planning to grow maize in 2017, and 40 who lived outside the four counties in which recruitment took place. Approximately 1,017 farmers lived in wards for which there was no soil test data available.

Table E3 reports summary statistics and balance checks and statistics for a range of different variables for each treatment. Baseline characteristics are balanced across treatment groups, with the exception of previous yields and land size, which are higher for the control group. We control for these characteristics in the main specifications, but results are robust to their exclusion.

During the 2017 long rains season, IPA/PAD sent messages to farmers in the treated groups encouraging them to experiment with locally appropriate quantities of agricultural lime and chemical fertilizers on a small portion of their farm. All the messages were based on ward-level soil test data (additional information about recommendations is presented in appendix B).[38] The messages focused on three types of recommendations: the use of agricultural lime in wards with median soil pH below 5.5, the use of planting fertilizer, and the use of topdressing fertilizers.

The SMS-based information service was shorter than the first IPA/PAD program and consisted of one welcome message followed by two sets of messages containing agronomic recommendations, each repeated twice. The complete list of messages is in appendix B. Messages were sent in either English or in Swahili, depending on farmers' language preferences at the time of registration. Farmers who lived in wards with pH measured to be lower than 5.5 received the following message: *"The soil in your area is [very] acidic. To avoid low yields treat now. Apply [quantity] bottletops of lime per planting hole. [quantity] kg for 1/4 acre"*. Farmers who lived in wards with pH higher than 5.5 received the following message: *"The soil in your area is slightly acidic. According to soil analysis, farms in your area do not require lime."* Farmers were also advised to use DAP at planting and urea at topdressing, contingent on rainfall realization. In particular, messages suggested to use urea if the rains were 'good' and use CAN otherwise.[39]

A random subset of farmers also received a phone call (or an SMS offer to receive a call) after each set of messages explaining the content of the text messages. This 15-minute phone call did not provide any additional information, but it allowed farmers to ask clarification questions to

[38]The information was at the ward level. A ward is an administrative unit in Kenya. Wards were chosen because they are one of the smallest units that farmers can self-report and that soil tests could be mapped into. In western Kenya, the average size of a ward is 12 km^2.

[39]Since it was not possible to have local rainfall patterns and make recommendations accordingly, farmers were provided with this information order to decide which fertilizer was more appropriate based on their own observation of the rains.

a PAD field officer and to hear the explanation multiple times. The purpose of the phone call was to strengthen the information provided via SMS.

The research team collected two types of outcome data. As in PAD/IPA1-K, all farmers participating in the experiment received two electronic coupons via SMS. Each coupon allowed farmers to obtain discounts on agricultural inputs from a local agricultural supply dealer. The first electronic coupon was redeemable for 15% on the first seven 10-kg bags of agricultural lime, and the second coupon provided a 15% discount on the first 1,000 Ksh (approximately $10 USD) spent on topdressing fertilizers (urea, CAN, or Mavuno).To avoid priming farmers about specific inputs, they were just told that the coupon would provide them with a discount for a range of agricultural inputs.[40] In addition, around 84% of farmers completed the endline survey with questions about their agricultural knowledge and input use during the season. We do not find evidence of differential attrition by treatment arm (appendix table E7, panel C, column 1).[41]

A.3 One Acre Fund's Programs

OAF is s across six countries in Eastern and Southern Africa. In 2017, they reported working with over 600,000 farmers (OAF, 2017). The OAF model relies on training farmers on modern agricultural techniques and providing them with seeds and fertilizer on credit. To receive the OAF input loan and training program, farmers must join a village group that is supported by a local OAF field officer. Farmers sign contracts with their field officers well before the agricultural season starts and get inputs delivered right before the beginning of the planting season. Farmers repay their loans at any time during the growing season. OAF clients form groups of eight to eleven farmers who participate in the program together through several shared activities, including signing a contract together and being jointly liable for their loans.

[40] To ensure that all farmers in treatment and control group were equally aware of the coupon, all farmers received a phone call a week before the program started, in which an enumerator explained how to use the coupon and at which agricultural supply dealers the coupons could be redeemed. 93% of farmers were reached during this activity.

[41] The survey was conducted in two batches: the first batch (approximately 40% of sample, randomly selected) was surveyed towards the end of the 2017 long rain season, while the second batch was implemented about 5 months later during the following short rains season, allowing for the collection of additional information on second season practices.

The standard bundle that OAF offers includes hybrid seeds and chemical fertilizers. However, to address the problem of high soil acidity, OAF started offering farmers agricultural lime as an optional add-on. Yet, across their many locations, demand for lime remained very low. Hypothesizing that this could reflect a lack of awareness, OAF designed and evaluated several informational programs to increase lime take-up. Since OAF field officers already follow detailed protocols, a key objective of the approach was to test cheap programs that would not require additional field officer training and delivery. We describe OAF's different strategies below.

A.3.1 Program 1 Kenya (OAF1-K)

Prior to 2016, less than 3% of OAF clients in western Kenya purchased agricultural lime through the organization (OAF, 2015). To increase take-up, OAF designed a phone-based extension pilot that consisted of six text messages targeting clients who had signed up for the OAF package during the previous season in a selected district of western Kenya.

OAF tested two versions of the messages. One group of farmers received simple SMS messages encouraging lime use and providing them with a customer engagement toll-free line which they could call to receive more information. The message read *"Hello [name], Your soil is acidic. Use lime to reduce acidity and increase yields. Call xxx-xxxx"*. A second randomly selected group of farmers received a more detailed message that mentioned the level of acidity measured in the farmer's area as well as the amount of lime recommended and expected return to its application: *"Hello [name], Your soil is [highly/moderately] acidic. We recommend [amount] kg of lime per acre at [total cost] Ksh. Use lime to reduce acidity and increase yields by [percentage]%.Call xxx-xxxx"*.[42] Customized messages were based on soil tests that had been previously conducted in the region. We discuss how these recommendations were constructed in appendix B. In total, 4,884 farmers participated, with 3,325 farmers randomly assigned to receive messages, and 1,559 farmers remaining as a control. The same SMS message was sent six times between August and September 2016, before the OAF input contract signing period, when farmers had to decide

[42] The percentage increase in yields depended on the local level of pH and the return estimated for that pH level based on OAF farm trials.

whether to request inputs from OAF for the following season.

A full list of balance checks in appendix table E4. Since OAF does not collect extensive demographic data we can only show a limited number of farmer characteristics at baseline. Running balance tests for twelve characteristics that OAF had for the farmers, which mostly included the products that farmers had purchased in previous seasons, we only find small differences at baseline: those in the treatment arms were more likely to purchase onion seeds and additional CAN fertilizer, and to receive a repayment incentive the previous year. We control for these variables in our main specifications, but the results are robust to their exclusion.

For OAF1-K we can measure outcomes using two sources of data: OAF administrative data and phone survey data collected by researchers. The administrative data contains information on loan enrollment and inputs purchased through the OAF program for the 2017 and 2018 long rain seasons. However, only 60% of farmers who received the text messages went to receive OAF loans in the 2017 long rain agricultural season. While we do not find evidence of differential participation in OAF programs by treatment arm (table E7, panel D, column 2), we take a conservative approach in our main specifications and define the outcome variable as lime purchased through OAF. This outcome is an imperfect measure of the overall effects of the program on lime purchases if farmers acquired lime from other sources. To explore this possibility and obtain additional information from farmers, a follow-up phone survey led by IPA was conducted in May 2017 with a random sample of 30% of the farmers participating in the trial. This survey asked respondents about their knowledge of lime and their input use during the 2017 long rains season. About 79% of selected farmers were surveyed, and we do not find differential treatment attrition for this sample (table E7, panel D, column 1). In September 2017, at the time of enrollment for the 2018 long rains season, a subset of the farmers who purchased inputs from OAF for the 2017 long rains season received additional messages encouraging lime adoption. The treatment assignment for this program was stratified on previous treatment status.

A.3.2 Program 2 Kenya (OAF2-K)

A second OAF program was implemented with approximately 30,000 farmers in four Kenyan districts in September 2017. Former OAF clients were randomized into a no message control group or a treatment group receiving SMS messages encouraging lime adoption (which did not depend on results from soil tests in the area). Additionally, a quarter of farmers were randomly assigned to receive additional messages encouraging the use of additional fertilizer (Extra CAN) for a second round of topdressing.

The messages randomly varied how the lime information was presented, number of repetitions (1 to 5 messages), and time between repetitions (every 2 to 8 days). Six main categories of messages were sent ranging from a basic messages that simply recommended to buy lime *"[Name], OAF recommends you register to buy Lime for your maize."*, to messages explicitly encouraging experimentation *[Name], OAF recommends you register to buy Lime for your maize. Try it on just a small part of your land to so that you and your neighbors can see the benefits."*, or leveraging on social comparison *" [Name], OAF recommends you register to buy Lime for your maize. Farmers all over Western are getting bigger yields by using lime. Keep up with them!"*. For simplicity, we pool all the different treatment arms in the main tables, but all the information on different treatment arms can be found in appendix B.[43]

Summary statistics and balance checks for treated and control farmers are reported in table E5. Apart from a small differences in land reported, we do not find evidence of consistent statistically significant differences at baseline between treatment and control arms (we also reject the null of joint significance). Farmers were later matched to OAF administrative data to measure their likelihood of demanding agricultural lime and other inputs for the following

[43] In the same period, OAF also conducted two other programs to encourage lime adoption via SMS. In Nambale district, where the OAF1-K program took place, a randomly selected subset of farmers that did not purchase lime during the previous season were matched to lime users and encouraged to talk to them to learn more about the product. A version of the OAF2-K program that did not involve topdressing fertilizer messages was implemented during the same period in Nambale. To simplify the exposition we exclude this district from the sample analyzed in this paper, however, its inclusion does not change the main results. Farmers outside the trial districts (excluding those in non-acidic areas) received different variations of SMS messages encouraging lime adoption. The content of the messages, number of repetitions, and frequency were randomly assigned. This component involved approximately 180,000 farmers. In this paper we focus exclusively on the results of the first program.

two agricultural seasons. Only 76% of farmers who received text messages decided to acquire any inputs through OAF, but we do not find differential likelihood of purchasing inputs by treatment status (E7, panel E). Again, we define the primary outcome variable as the probability of purchasing agricultural lime from OAF.

In September 2018, at the time of enrollment for the 2019 long rains season, all the farmers who purchased inputs from OAF for the 2018 long rains season received additional messages encouraging lime adoption (but no messages about fertilizer).

A.3.3 Program 3 Rwanda (OAF3-R)

In 2017 a modified version of the Kenya program was implemented in Rwanda. In Rwanda, OAF partners with the government to provide goods, services, and training to rural farmers. Since 2016, OAF and the government of Rwanda have engaged in a concerted effort to promote adoption of travertine, a type of agricultural lime. Activities involved marketing lime, widespread soil pH testing, and offering substantial price subsidies (75% off the price) in several districts. OAF reported that in districts where the price subsidy was offered, lime demand went up from 7% to 21%. Since all these interventions were costly, OAF also decided to test the effectiveness of text-messages as an inexpensive way to increase lime use.

Unlike the Kenya OAF program that operates only during the main agricultural season, the OAF Rwanda provides inputs on credit for both seasons. Farmers who want to purchase inputs for the secondary season (February to August) need to place their orders before the beginning of the previous main agricultural season (September to January), but are allowed to drop some products before the time of delivery.

In June 2017, during the enrollment period for the 2018 main agricultural season (September 2017 to January 2018), a large-scale program aimed at increasing lime adoption through the use SMS messages was implemented in all districts where OAF operates. Since phone ownership is much lower in Rwanda than in Kenya (only 53% of the farmers registered in the program had a phone number reported in the database) one of the objective of this program was to measure spillovers among farmers in the same group, in particular to those who did not own a phone. Therefore, the randomization was done at the farmer group level, with some groups partially treated. Since the focus of this paper is to estimate the direct effect of the program, we exclude from the main analysis the farmers in the partially treated groups that were assigned not to receive messages. In presence of spillovers within farmer groups, their inclusion would lead to underestimating the treatment effect, as discussed in section 6. A full description of this program and experimental design can be found in Harigaya et al. (2018).

As in OAF2-K, the messages varied content, framing, and number of repetitions.[44] Seven types of messages were sent, ranging from a general promotions *"Many fields in Rwanda have acidic soil and need LIME to increase yields. Order from OAF now."*, to messages explaining that *"[acidity] blocks fertilizer uptake."* and *"Applying LIME solves the problem, increasing crop yields"*. Other messages tried to create a sense of urgency by using wording like *"Order it immediately"*. All messages were either gain-framed or loss-frame with respect to yield increases generated by lime use. In addition to these messages, farmers in half of the treated groups received an additional message encouraging them to share the information with other farmers, especially those without phone. Additional information, including the complete list of messages, can be found in appendix B.

From a total of 216,475 farmers registered in the OAF program, only 114,569 had a phone registered in the database, and 85,160 had a unique phone number. Since the unit of randomization was the group and farmer (rather than phone number), some phone numbers shared among more than one farmer were sometimes assigned to multiple treatments. In our analysis we drop all farmers that did not have a phone registered in the database and consider the original treatment assignment, regardless of whether phones are shared or not. The main results are robust to excluding all farmers with shared phones from the analysis. Summary statistics and balance checks are in table E6. We detect some small differences in group characteristics and input purchases from OAF in the 2017 main agricultural season but we do not reject the null of joint significance. We can measure whether farmers purchased lime from OAF for the 2018

[44] In Rwanda, OAF is known as Tubura and offers a type of lime known as travertine. The messages read "TRAVERTINE" and "TUBURA", not "LIME" and "OAF". We made these changes to simplify the exposition.

agricultural seasons. Table E7 panel F shows that only 65% of control farmers enrolled in the OAF loan program for the main agricultural season, but we do not find differential likelihood of the decision to purchase any inputs by treatment status.

B Message Content

In this section we provide information on message content and explain how farmers were provided with local information.

B.1 KALRO

KALRO's e-extension program consisted in 20 SMS messages send in the period corresponding to the 2015 short rains season: June-November 2015. The first set of messages were in English. Mid-intervention the messages were switched to Swahili.[45] We list all messages sent by KALRO below:

- We at KALRO- Kakamega shall be sending you 20 SMS tips on how to increase your maize and legume (beans, groundnuts, soybeans) yield
- Keep all the records of your farming activities including inputs and outputs to help you know whether your farming is profitable
- Test your soil after every 4 years. Enquiries: KALRO Tel:[phone] or Soil Cares Ltd: [phone]
- If soil is acidic (pH less than 5.5), apply recommended rate of agricultural lime at least 30 days before planting. Enquiries: Tel.[phone]
- Construct raised bands and trenches to control soil erosion, reduce nutrient loss and keep rain water in the soil
- Add and/or leave all organic matter (manure, crop/weed residues and compost) to your field. Do not burn your fields. Burning destroys useful microorganisms.
- Prepare land early, at least one plough and one harrow, ready for planting before onset of rains
- Plant before or at the onset of rains. Plant on well drained, fertile soils
- Use certified maize and legume seed recommended for your area, bought from an approved agro-dealer. Use 10 kg maize seed and 40kg of legume seed per acre. Enquiries: [phone]
- Maize and legumes planted in rows are easier to weed & apply fertilizer. You may plant maize alone/pure or together with legumes as follows:
- For pure maize make rows 2.5 feet (75cm) apart and holes 1 foot (30cm) apart along the row. Place 2 and 1 maize seeds in alternate holes.
- For maize and legume intercrop, plant maize as for pure stand and one row of legume (beans, soybean or groundnut) between two maize rows at spacing of 10cm from one hole to another.
- For better maize and legume harvests, inoculate legumes, rotate or intercrop, use fertilizer and manage your crop and soils appropriately.
- Use fertilizer to increases yields. Apply 1 heaped Fanta top of NPK or DAP in each hole for maize, cover with little soil, add seed and cover seed with soil. Fertilizer MUST not touch the seed
- Weeds compete with your crops for nutrients and so reduce yields. Keep fields free of weeds and pests. Thin maize seedlings to 1 plant per hole as you weed.
- Topdress your maize with a level Fanta bottle top of CAN or Mavuno top dress fertilizers 6 weeks after planting. Apply around each plant-5cm away and cover with soil. Apply when soil is moist.
- Harvest as soon as the crops are mature. For maize look for the black eye; for legumes when 90-100% of pods are brown. In late harvests, termites, rodents, insects, diseases birds reduce yield.
- Remove husk from maize cobs in the field to avoid transporting weevils from the field to the store. The husks will improve the organic matter in the soil.
- Dry your harvest in open sun, but protect it from rain. Thresh/shell and redry to moisture content of 11-12%.

[45] While 75% of farmers report speaking English at baseline, there is a risk that some farmers might have not understood the initial messages. We do not find heterogeneous treatment effects by language spoken.

B.2 IPA/PAD1-K

The first program implemented by IPA and PAD consisted in a series of 24-28 messages sent during the 2016 short rains season: August-December 2016. Two versions of the service were tested. The first, denoted as "General" provided blanket recommendations on maize farming in western Kenya. The second, denoted as "Specific", included customized recommendations for planting fertilizer and agricultural lime based on local soil characteristics.

Farmers participating in this programs were recruited from two sources: a database of farmers who had previously participated in IPA activities (IPA farmers), and administrative record of Mumias Sugar Company, a company that works with contract farmers in the area (MSC farmers). In order to construct customized recommendations for the specific messages, farmers were linked to a local landmark that could then be matched with soil data.This is a context in which there are no addresses and a lot of variation on how village names are reported. Therefore it was difficult for farmers to report their exact location. Primary schools are often used as landmarks. IPA farmers were matched to the primary school (usually the closest one to their farm) and provided recommendation based on median soil characteristics (exchangeable acidity and phosphorous) obtained from soil tests performed in the 2 km area around the school. The soil data were collected for previous projects by IPA (Fabregas et al., 2017b) and analyzed by the Kenya Agricultural Research Institute (KARI) using wet chemistry in 2011 and 2014. MSC farmers were matched to their "field", a set of plots cultivated by multiple farmers and aggregated by the company for organizing their activity, including soil testing. The recommendations provided to them were based on median soil characteristics (pH and phosphorus) of the sample collected from that field and analyzed by MSC in the period 2009-2016.

Since the topdressing fertilizer recommendation were not specific to the farmers' catchment area, but based on the quantity of nitrogen required to achieve a certain expected yield, specific application rates were provided to all treated farmers. Messages were sent either in English or in Swahili, depending on farmers' preferences indicated during the baseline phone survey. We report all the messages below: [G] indicates that the message was received by the General treatment group, and [S] denotes it was received by the Specific treatment group.

- [G/S]: Welcome to PAD's SMS information service. We will give you tips on agricultural inputs to apply on 1/8 of an acre so you can experiment during this short rains season. Receiving SMS messages is free.
- [G]: High soil acidity levels reduce nutrients available to plants, such as phosphorus, which causes symptoms of stunted growth and purple colouration of maize.
- [S]: Previous soil tests of shambas around [landmark] showed [degree] soil acidity levels. High acidity levels reduce nutrients available to plants, such as phosphorus, which causes symptoms of stunted growth and purple colouration of maize.
- [G]: Lime reduces soil acidity and makes nutrients such as phosphorus available for your maize.
- [S]: Based on soil tests of shambas around [landmark], we recommend you buy [quantity] kg of lime, [quantity] kg of DAP, and 6 kg of urea for microdosing 1/8 acre of your maize. Lime reduces soil acidity and makes phosphorus available for your maize.
- [S]: We would like you to try our recommendations in 1/8 of an acre. To measure 1/8 of an acre, walk around your farm and draw a square with each side 33 steps long. Walk normally, don't make long strides. If you land is a rectangle, the sum of 2 sides should measure in total 66 steps. Start from a corner, walk along the short side, count your steps until you reach the end. Turn around and keep walking along the long side until you finish counting 66 steps.
- [S]: When planting this season try adding a layer of lime [quantity] bottletop, then cover with soil and add a second layer of DAP ([quantity] bottletop) per hole on 1/8 acre to correct soil acidity and make more nutrients available for your plants. Apply 1 bottletop of urea per hole at top dressing.
- [G]: Use a ruler or measured rope to plant maize in rows using correct spacing of 75 cm x 25 cm. This offers maximum yield while limiting competition for nutrients, light and water.
- [S]: Use a ruler or measured rope to plant maize in rows using correct spacing of 75 cm x 25 cm. This offers maximum yield while limiting competition for nutrients, light, and water. You should be able to fit 2580 planting holes in 1/8 of an acre. Use sisal twine to encircle this area so you can compare the results at harvest.
- [S]: Have you bought lime and DAP yet? If not, buy a total of [quantity] kg of lime and use with [quantity] kg DAP for microdosing on 1/8 of your acre. DAP is the most cost efficient source of phosphorous. When lime is combined with DAP, it reduces soil acidity and makes nutrients available for your maize.
- [G]: Calcium lime is safer for your health and the plant. This lime could be either brown or grey.
- [S]: [agrovet] will be stocked with lime (calcium lime) and DAP during this short rain season. This lime is brown and it is safer for your health and the plant. It is also heavier than the white lime so you only need to apply [quantity] bottletop per plant. The price of lime today is Ksh 7 per kg. The price of DAP today is Ksh [price] per kg.
- [G/S]: Plant maize seed when there is enough moisture after 2-3 rains, to enable absorption of water by seed and fertilizer. Delayed planting leads to reduced yields. To stop receiving these SMS messages reply "STOP".
- [G/S]: Plant two maize seeds per hole to ensure one survives. Do not use broken or damaged seeds because they will not germinate. Use certified seeds, they grow faster and are high yielding.
- [G]: Are you ready to plant your maize? We recommend you apply both lime and fertilizer in micro-doses at planting. 5 weeks later we recommend you apply top dressing fertilizer in micro-doses
- [S]: Do you know the 5 Golden Rules for successful micro-dosing? Based on soil tests performed around [landmark], we recommend you to: Apply [quantity] bottletop of lime and cover with soil and then add [quantity] bottletop of DAP. Cover with 2 inches of soil. Use 2 seeds per planting hole.Cover the seeds with 2 inches of loose soil. Apply 1 bottletop of urea as top dressing fertilizer 5 weeks later when the plant is knee high.
- [G/S]: Remember, lime should only be used during planting and not at top dressing. Lime is not a fertilizer and could burn the plant if applied at top dressing.
- [G/S]: At planting, if you are applying lime in micro-doses, remember to cover it with soil before applying fertilizer and planting seeds. Lime should not be in direct contact with the seeds as it may burn them. When you apply lime, wear protective clothing such as long sleeves and gloves. Cover your mouth and nose with a scarf and wear goggles.
- [G/S]: Gap your maize immediately after emergence. Gapping is done by re-planting maize seeds in places that have not germinated. This gives you optimum plant population that leads to optimum yields.
- [G/S]: During first weeding, thin to one maize plant per hole. You should remove striga immediately to reduce competition for nutrients and water, and to prevent stunted growth!
- [G]: Have you already planted your maize this season? If not, we recommend applying lime at planting. Lime reduces soil acidity and makes nutrients such as phosphorus available for your maize.
- [S]: Have you already planted your maize this season? If not, we recommend applying lime at planting. We recommend you apply [quantity] bottletop per planting hole. Buy [quantity] kg of lime to experiment on 1/8 of an acre. Lime reduces soil acidity and makes nutrients such as phosphorus available for your maize.
- [G]: If you applied lime on your maize at planting, we recommend using urea at top dressing because it is a less expensive source of nitrogen.
- [S]: If you applied lime on your maize at planting, we recommend using urea for top dressing because it is a less expensive source of nitrogen. Buy 6 kg of urea for use on 1/8 of an acre.
- [S]: [agrovet] will be stocked with urea during this short rain season. The price of urea is Ksh [agrovet] per kg.
- [G]: When the maize reaches knee high (5 weeks after planting), apply top dressing fertilizer.
- [S]: When the maize reaches knee high(5 weeks after planting), based on soil tests around **[landmark]**, we recommend you apply 1/2 bottletop of urea per plant, making a 15 cm circle around the maize plant.
- [G/S]:Conduct second weeding 6 or 7 weeks after planting. Uproot all striga before it produces seeds because it reduces maize yields if not removed
- [G/S]: We invite you to participate in an SMS poll to help you recognize potential maize diseases and provide advice. Reply OK to start. Messages are free.
  - Do you see straight lines of holes on newly formed maize leaves?
    [if yes] This could be stalk borers. Apply insecticide e.g. bulldock or tremor, into the funnel or spay the maize plant with pentagon at top dressing.We hope this information was helpful. We will be sending another poll question tomorrow. Thank you!
    [if no] This is good news! Thank you for answering our question. We will send another question tomorrow.
  - Do you notice yellow or white streaks or discoloration on the leaves of your stunted maize plants? [if yes] It could be Maize Streak Virus. Eradicate grass weeds and use malathion or dimethoate to control as soon as possible. We hope this information was helpful. We will be sending another poll question tomorrow. Thank you!
    [if no] This is good news! Thank you for answering our question. We will send another question tomorrow.
  - Do you see striga weed in your maize plot? Striga has thin leaves and pink or purple flowers and attaches onto the maize roots.
    [if yes] Uproot all striga that has emerged. Striga competes with your maize for nutrients, water, and light and leads to reduced maize yields. We hope this information was helpful. We will be sending another poll question tomorrow. Thank you!
    [if no] This is good news! Thank you for answering our question. We will send another question tomorrow.
  - Do you see ants that cut maize stalks and feed on fallen maize cobs?
    [if yes] It could be termites.

Dig out all anthills around your maize farm and ensure that you destroy the queen. Alternatively, you can dig a deep hole at the center of the anthill and use insecticide to kill the ants. We hope this information was helpful. This is the last poll question. We will NOT send another question tomorrow. Thank you for your participation!
[if no] This is good news! This is the last poll question. We will NOT send another question tomorrow. Thank you for your participation!

- [G/S]: WEEDING REMINDER! Conduct second weeding 6 or 7 weeks after planting. Weeds compete with your maize for nutrients, water, and light, which reduces yields.
- [G]: Have you already applied top dressing fertilizer on your maize? If not, we recommend using urea at top dressing because it is a less expensive source of nitrogen.
- [S]:Have you already applied top dressing fertilizer on your maize? If not, we recommend using urea at top dressing because it is a less expensive source of nitrogen. Buy 6 kg of urea for use on 1/8 of an acre and apply 1/2 bottletop of urea per plant. Apply urea when there is enough moisture in the soil to avoid loss through evaporation.
- [G/S]: Harvest maize at physiological maturity when cobs droop and leaves dry. Dry maize in the sun even after shelling to avoid mold and attack by weevils. Maize grain must remain dry and clean during storage to avoid reduction in quantity and quality.
- [G/S]: We hope you enjoyed these messages from Precision Agriculture for Development. Our team will follow up with a phone call in the coming weeks to hear more about how your planting season went.

During the 2017 long rains season, all treated farmers received 5 identical SMS messages about agricultural lime. These messages were based on the local recommendations constructed for the IPA/PAD2-K program. We report the text these messages below:

- [If pH≤5.5]: The soil in your area is [level] acidic. To avoid low yields, treat now. Apply [quantity] bottletop of lime per planting hole. [quantity] lime per 1/4 acre.
- [If pH>5.5]: The soil in your area is slightly acidic. According to our analysis, farms in your area do not need lime.

B.3 IPA/PAD2-K

The second program implemented by IPA and PAD consisted in 3 message about planting inputs for maize farmers (lime and fertilizer), repeated twice, plus 2 additional messages on topdressing fertilizer, also repeated twice. Planting recommendations were based on local soil data: ward level median level of pH and phosphorous, and target yield of 2 t/ha, while topdressing recommendations were only based on target yield.[46]

Recommendations for lime and DAP were provided based on median soil characteristics in the farmers' ward.[47] The soil data used to generate these recommendations was obtained by pooling data collected by 4 different organizations: IPA, OAF, Mumias Sugar Company, and

[46]The target yield of 2 t/ha aimed at generating an improvement over the baseline average of 1.42 t/ha, while keeping the cost of the input package affordable for farmers. The government's recommended application of phosphorus for western Kenyan soils, for a target yield of 3.9 t/ha in soils with P below 10 mg/kg, is 26kg P/ha, corresponding to 130 kg DAP/ha, (FURP, 1995; Wasonga et al., 2008). With a target yield of 2 t/ha, the recommendations provided as part of this program involved applying 21kg P/ha, corresponding to 107 kg DAP/ha.

[47]Recommendations were provided based at the ward level because that is the most precise information collected about farmers' location. The data was aggregated into medians because the majority of the soil data available was not geocoded and only provides information on the administrative unit in which the sample was collected.

the German Agro Action (Welthungerhilfe).[48] These sources provided over 30,000 soil tests for program area. However, in order to base the recommendations on the most recent data, data was dropped for soil tests performed before 2014, when possible. The final dataset used included about 7,085 observations for 108 wards.[49]

Messages were sent either in English or in Swahili, depending on farmers' preferences indicated during the baseline phone survey. We list the messages below:

- Welcome to PAD, IPA's free advice service for maize growers. You will receive advice for your needs based on more than 10,000 soil tests from Western Kenya.
- The soil in your area is [level] acidic. To avoid low yields treat now. Apply [quantity] bottle top of lime per planting hole. [quantity] kgs for 1/4 acre. OR The soil in your area is slightly acidic. According to our soil analysis, farms in your area do not need lime.
- Soil acidity causes stunted growth.Lime reduces soil acidity and makes nutrients of DAP more available for your maize.
- When planting, apply [quantity] bottle top of lime. Cover with a handful of soil. Add [quantity] bottletop of DAP, cover with enough soil to avoid direct contact of inputs. OR When planting, apply [quantity] bottle top of DAP, cover withe enough soil to avoid direct contact of inputs.
- Check your phone! We sent you 3 planting recommendations last week [If you flash [number] before Friday this week, we will you callback soon to explain them/We will call you soon to explain them]
- Top-dress when your maize has more than 4 leaves up to knee high. If rains are good.apply 3/4 bottle top of UREA. If rains are low, apply 3/4 bottle top of CAN.
- UREA can increase your maize yields as much as CAN if rains are good. Try 11 kg of urea in 1/4 acre and see the results
- Check your phone! We sent you 2 top-dressing messages this week [If you reply YES or flash [phone] by Tuesday, we will call you back soon to explain them/We will call you soon to explain them.]

B.4 OAF1-K

In September 2016, during the period in which OAF farmers were placing their orders for the 2017 long rains season, OAF sent SMS messages about soil acidity and agricultural lime. Two types of messages were sent: the first, denoted as "Broad", simply encouraged farmers to use lime to reduce soil acidity and increase yields, while the second, denoted as "Detailed" provided recommendations on lime application rates and expected yield increase customized to the farmers' site.

Although the standard application rate recommended by OAF and reflected in field mate-

[48]The IPA dataset was assembled in 2011 and 2014 in Busia county for previous projects (Fabregas et al., 2017a) and extended in 2016 as part of test plot activities in the same area. The OAF data was collected in 2016 across the entire study area. Mumias Sugar Company shared the data they collected for their operations in Busia and Kakamega counties between 2009 and 2016. The German Agro Action data was collected in Kakamega and Siaya counties in 2015.

[49]Data collected before 2014 was dropped if at least 30 more recent observations in the ward were available. Since the data displays clear trends of decreasing pH and phosphorus levels over time, they were adjusted using coefficients based on the Mumias Sugar Company soil data: a coefficient of -0.027 per year was applied for pH and -0.504 per year for phosphorus. These coefficients were obtained by regressing pH and phosphorus data on a time trend and constant, controlling for field fixed effects, these regressions are based on a sample of over 60,000 observations.

rials was 200kg/acre across the entire program, the detailed message encourages farmers to use different application rates based on the pH level predicted for the farmers' site.[50] To obtain these predictions, OAF used their own own soil tests, performed using soil spectroscopy, and soil data collected for a previous project by IPA (Fabregas et al., 2017b). These soil chemistry results were then interpolated across areas through Kriging to create a continuous field of soil chemistry predictions. Optimal lime application rates, for each level of pH, were based on OAF on-farm agronomic trials conducted in 2015 (OAF, 2015). During that trial three different lime application rates were tested: 50kg/acre, 100kg/acre and 200kg/acre. The sample was divided according to pH quintiles and, for each quintile, the lime application rates that resulted in the most precisely estimated effect on yield was chosen. Two different lime application rates were recommended, based on the local predicted level of pH: 200kg/acre and 50kg/acre.[51]

Farmers in both treatment groups received 6 identical messages, all messages were sent in Swahili. We report the messages below:

- [Broad]: Hello [name],Your soil is acidic. Use lime to reduce acidity and increase yields.Call xxx-xxxx.

- [Detailed]: Hello [name],Your soil is [level] acidic. We recommend [amount] kg of LIME per acre at [total cost] Ksh. Use lime to reduce acidity and increase yields [percentage increase]%.Call xxx-xxx.

B.5 OAF2-K

In September 2017, when OAF farmers were enrolling for the 2018 long rains season, OAF implemented a second program aimed at encouraging lime adoption. In addition subset of farmers was randomly assigned to receiving additional messages encouraging use of an extra amount of topdressing fertilizer (Extra CAN).

Six different types of messages were sent: a "Basic" message simply recommended to purchase lime, a message, "Yield increase", also mentioned that lime would increase yields, two encouraged experimentation, "Experimentation (selfish)" and "Experimentation (neighbors)", and two leveraged on behavioral nudges "Social comparison" and "Self-efficacy". Half of the

[50]Since OAF does not collect the coordinates of farmers' plots, farmers were assigned to the GPS coordinates of the site to which inputs are delivered by OAF.

[51]Robert On, Matthew Lowes, and David Guerena produced these recommendations.

treated farmers were randomly assigned to receive messages addressing the whole family instead of the individual (by replacing the word "you" with "your family"). The messages encouraging use of additional quantities were identical to those encouraging use of lime (the word "Lime" was replaced by "Extra CAN"). Farmers assigned to receive both lime and fertilizer message were randomly assigned to receive one of the two first and the other on the next day for all repetitions. The number of repetitions (from 1 to 5) and the frequency of the messages (every 2, 4, 6, or 8 days) were cross-randomized.

We report all messages below:

- [M1: Basic] [Name], OAF recommends [you/your family] register to buy [Lime/Extra CAN] for your maize.
- [M2: Yield increase] [Name], OAF recommends [you/your family] register to buy [Lime/Extra CAN] for your maize. You'll get higher yields by using [Lime/Extra CAN].
- [M3: Experimentation (selfish)] [Name], OAF recommends [you/your family] register to buy [Lime/Extra CAN] for your maize. Try it on just a small part of your land to see the benefits.
- [M4: Experimentation (neighbors)] [Name], OAF recommends [you/your family] register to buy [Lime/Extra CAN] for your maize. Try it on just a small part of your land to so that you and your neighbors can see the benefits.
- [M5: Social Comparison] [Name], OAF recommends [you/your family] register to buy [Lime/Extra CAN] for your maize. Farmers all over Western are getting bigger yields by using [Lime/Extra CAN]. Keep up with them!
- [M6: Self-efficacy] [Name], OAF recommends [you/your family] register to buy [Lime/Extra CAN] for your maize. You have the ability to achieve higher yields by using [Lime/Extra CAN]!

B.6 OAF3-R

OAF-Rwanda, known as Tubura, implemented an SMS-based program aimed at encouraging experimentation with a type of agricultural lime, known in Rwanda as travertine. The messages were sent in June 2017, when farmers were enrolling for the 2018 main agricultural season (September 2017 to January 2018) and the following secondary agricultural season (February to August 2018). As for OAF2-K, The purpose of this program was to understand how to optimize message content, framing, number of repetitions and framing. In addition, given the relatively low mobile phone penetration in the country, OAF wanted to explore ways to increase spillovers within farmers' group in order to reach all farmers. For this reason, the first stage of randomization took place at the group level, assigning farmers groups to four group-level treatments: a pure control group where no farmers in the group (GO: No SMS), a treatment in which all farmers received identical messages (G1: Same SMS), a treatment in which all farmers in the group received messages, but content and framing were randomly assigned at the individual level (G2: Diff SMS), and a treatment in which farmers received messages with probability 0.5 and content and framing were randomized at the individual level (G3). In this paper we focus on the direct effect of the receiving messages on individual farmer rather than on group level outcomes. Therefore, we divide farmers in G3 into messages receiving (G3 - Treated) and non message receiving (G3 - Control). We exclude from our analysis all farmers who did not have a phone number registered in OAF's database.

Seven types of messages were sent: a basic message encouraged to purchase lime "General Promotion", the second indicated the application rate and expected impact "Specific + yield impact", the third helped farmers assess their need for lime "Self-diagnosis", the fourth encouraged farmers to have their soil tested "Soil test", the firth explained that lime can be used to increase fertilizer efficiency "How travertine works", the sixth encouraged farmers to order lime immediately "Order immediately", and the seventh indicated that acidity was a problem in the farmer's area "Your cell is acidic + yield impact". All messages were either framed positively (gain) or negatively (loss) and the number of repetitions varied (from 1 to 4). Finally, message receiving farmers in half of the treated groups received an additional message encouraging them to spread the information to others in their group, especially those without phones (Social nudge message).

- [T1-G: General promotion (gain)] Many fields in Rwanda have acidic soil and need TRAVERTINE to increase yields. Order from TUBURA now.
- [T1-L: General promotion (loss)] Many fields in Rwanda have acidic soil and need TRAVERTINE to avoid a yield loss. Order from TUBURA now.
- [T2-G: Specific+ yield impact (gain)] Many fields in Rwanda have acidic soil. Applying 25 kg/are of TRAVERTINE will increase yields by 20%.Order from TUBURA now.
- [T2-L: Specific+ yield impact (gain)] Many fields in Rwanda have acidic soil. Applying 25 kg/are of TRAVERTINE will prevent a yield loss of 20%. Order from TUBURA now.
- [T3-G: Self-diagnosis (gain)] Do you have fields with poor harvests even when you use fertilizer? You probably have acidity and need TRAVERTINE to increase yields. Order from TUBURA now.
- [T3-L: Self-diagnosis (loss)] Do you have fields with poor harvests even when you use fertilizer? You probably have acidity and need TRAVERTINE to avoid a yield loss. Order from TUBURA now.
- [T4-G: Soil test (gain)] Ask your Field Officer for a free soil test to learn if your fields are acidic and you need to order TRAVERTINE to increase yields.
- [T4-L: Soil test (loss)] Ask your Field Officer for a free soil test to learn if your fields are acidic and you need to order TRAVERTINE to avoid a yield loss.
- [T5-G: How travertine works (gain)] Many fields in Rwanda have acidity, which blocks fertilizer uptake. Applying TRAVERTINE solves the problem, increasing crop yields. Order from TUBURA now.
- [T5-L: How travertine works (loss)] Many fields in Rwanda have acidity, which blocks fertilizer uptake. Applying TRAVERTINE solves the problem, preventing a yield loss. Order from TUBURA now.
- [T6-G: Order immediately (gain)] Many fields in Rwanda have acidic soil and need TRAVERTINE to increase yields. Order it immediately, when signing your TUBURA order form.
- [T6-L: Order immediately (loss)] Many fields in Rwanda have acidic soil and need TRAVERTINE to avoid a yield loss. Order it immediately, when signing your TUBURA order form.
- [T7-G: Your cell is acidic + yield impact (gain)] In your cell the soil is acidic. If you apply 25 kg/are of TRAVERTINE you can boost yields by 20%. Order from TUBURA now.
- [T7-L: Your cell is acidic + yield impact (loss)] In your cell the soil is acidic. If you apply 25 kg/are of TRAVERTINE you can avoid a yield loss of 20]%. Order from TUBURA now.

Social nudge message:

- [SN] Please share this information about TRAVERTINE with your group members and neighbors, especially those who don't have phones!

C Local Agricultural Recommendations

In this section we briefly describe how the local agronomic recommendations were constructed.

IPA-K/PAD1-K Lime recommendation for IPA-K farmers where calculated based on median level of exchangeable acidity in the area. Since exchangeable acidity information was not available for the MSC sample, lime recommendations were based on the median level of pH in the farmers' fields.

The amount of planting fertilizer recommended was based on the median amount of phosphorus measured in the area which determined the recommended quantity of diammonium phosphate (DAP).

Topdressing fertilizer recommendations were based on the quantity of nitrogen required to achieve a certain expected yield. The quantity was selected based on the target yield of 2 t/ha. For this target yield the quantity of nitrogen required is 54 kg per hectare, which corresponds to 117 kg of urea or 206 kg of calcium ammonium nitrate (CAN). Given the lower amount required and the fact that the average price of urea was lower than that of CAN (66 Ksh vs 76 Ksh per kg) urea was recommended.

IPA-K/PAD2-K

Farmers participating in this study were provided with lime and planting fertilizer (DAP) recommendations based on the local level of soil acidity, measured in terms of soil pH, and phosphorus, respectively. The recommended input quantities were based on a target yield of 2 t/ha, which represents an improvement with respect to the baseline average of 1.42 t/ha, while keeping the cost of the input package affordable for the farmers.

The amount of lime recommended was decreasing in the level of pH while the amount of DAP recommended[52] was decreasing in the level of phosphorus.

The recommendations are based on micro-dosing, rather than general broadcasting methods,

[52] The government's recommended application of phosphorus for Western Kenyan soils, for a target yield of 3.9 t/ha in soils with P below 10 mg/kg, is 26kg P/ha, corresponding to 130 kg DAP/ha, (FURP, 1995; Wasonga et al., 2008). With a target yield of 2 t/ha, the recommendations provided as part of this study involved applying 21kg P/ha, corresponding to 107 kg DAP/ha.

to maximize effectiveness[53] (IPNI, 1999). To provide a standard measure for micro-dosing recommendations, farmers were advised to use a soda bottletop, which is a common item easily available throughout the study area.

Recommendations for lime and DAP were provided based on median soil characteristics in the farmers' ward.[54] Soil data was pooled from 5 different sources: (1) Soil data collected by IPA-K in Busia county for previous projects (Fabregas et al., 2017a) in 2011 and 2014 and as part of test plot activities conducted in 2016. (2) Soil data collected by One Acre Fund across the entire study area in 2016. (3) Soil data collected by Mumias Sugar Company in Busia and Kakamega counties between 2009 and 2016. (3) Soil data collected by the German Agro Action (Welthungerhilfe) in Kakamega and Siaya counties in 2015.

These datasets provided over 30,000 soil tests for the area in which the study took place, specifically, the set of wards in which the farmers participating in the intervention are based. However, in order to base the recommendations on the most recent data, data was dropped for soil tests performed before 2014, when possible[55] The final dataset used included about 7,085 observations for 108 wards.

One Acre Fund

The standard application rate recommended by OAF and reflected in field materials was 200kg/acre across the entire program. In order to generate "local" recommendations OAF's used their own own soil tests, performed using soil spectroscopy, and soil data collected for a previous project by IPA-K (Fabregas et al., 2017b) and analyzed by the Kenya Agricultural

[53]Farmers were also recommended to use micro-dosing for lime application as it requires lower investment and yields higher returns in the short term (Mortvedt and Follet, 1999; Terman and Engelstad, 1976; Plaster, 2003; OAF, 2015). This practice is not recommended by the local government and the Kenya Agriculture and Livestock Research Organization, which recommend broadcasting. However, it is in line with the recommendations of One Acre Fund, an NGO that serves about 400,000 farmers in the region, including 35% of the farmers in our sample.

[54]Recommendations were provided based at the ward level because that is the most precise information collected about farmers' location. The data was aggregated into medians because the majority of the soil data available was not geocoded and only provides information on the administrative unit in which the sample was collected.

[55]Data collected before 2014 was dropped if at least 30 more recent observations in the ward were available. Since the data displays clear trends of decreasing pH and phosphorus levels over time, they were adjusted using coefficients based on the Mumias Sugar Company soil data: a coefficient of -0.027 per year was applied for pH and -0.504 per year for phosphorus. These coefficients were obtained by regressing pH and phosphorus data on a time trend and constant, controlling for field fixed effects, these regressions are based on a sample of over 60,000 observations.

Research Institute (KARI) using wet chemistry in 2011 and 2014. These soil chemistry results were then interpolated across areas through Kriging to create a continuous field of soil chemistry predictions. Since OAF does not collect the coordinates of farmers' plots, farmers were assigned to the GPS coordinates of the site to which inputs are delivered by OAF.

Optimal lime application rates, for each level of pH, were based on OAF on-farm agronomic trials conducted in 2015 (OAF, 2015). During that trial three different lime application rates were tested: 50kg/acre, 100kg/acre and 200kg/acre. The sample was divided according to pH quintiles and, for each quintile, the lime application rates that resulted in the most precisely estimated effect on yield was chosen. Two different lime application rates were recommended, based on the local predicted level of pH: 200kg/acre and 50kg/acre.

D Regression Controls and Variables

- The KALRO sample includes controls for gender, hearing about lime at baseline, index of baseline input use, grown legumes, land size, baseline knowledge about soil tests.
- IPA/PAD1-K include age, gender, primary education, sample of origin, preferred language, phone network, farm size, knowledge score at baseline, previous input use, and measures of interest in the program at baseline.
- IPA/PAD2-K include: age, gender, preferred language, farm size, previous lime use, and agricultural supply dealer (recruiter) dummies.
- OAF1-K sample includes controls for number of seasons in the program, group size, repayment incentives received, ordering: size of maize package, bean seeds for inter-cropping, compost boost products, solar lamps, cook stoves, extra CAN, harvest sheets, storage bags, onion seeds, health insurance, and sanitary pads.
- OAF2-K sample includes controls for number of seasons in the program, group size, predicted pH level in the area, size of the maize package, and indicators for whether the farmer purchased solar lamps and extra CAN in the previous season.

- OAF3-R sample includes controls for number of seasons in the program, group size, and administrative information from 2017 main agricultural season including: credit size, quantity of planting fertilizer (DAP and NPK) purchased, and indicators of whether the farmer purchased lime and urea in 2017.

Table D1: Inputs

Sample	Recommended	Other
KALRO	Lime, planting fertilizer (DAP, NPK), topdressing fertilizer (CAN, Mavuno), compost, manure, hybrid seeds	Rhizobia, striga control, pest and disease control, storage bags
IPA/PAD1-K	Lime, DAP, urea	NPK, CAN, Mavuno
IPA/PAD2-K	Lime, DAP, urea	NPK, Mavuno, hybrid seeds, pesticides
OAF1-K	Lime	Actellic, compost, extra CAN, drying sheets, storage bags, machete, hoe
OAF2-K	Lime, extra CAN	Actellic, compost, drying sheets, storage bags
OAF3-R	Lime	DAP, NPK, urea, storage bags

E Attrition & Balance

Table E1: KALRO:Summary Statistics & Balance

	Control (1)	Treated (2)	(1) vs. (2) (3)
Age	41.29	39.79	1.50
	(0.66)	(0.65)	(0.92)
Female	0.65	0.65	-0.01
	(0.02)	(0.02)	(0.03)
Primary school	0.53	0.54	-0.01
	(0.02)	(0.02)	(0.03)
Secondary school	0.03	0.04	-0.01
	(0.01)	(0.01)	(0.01)
Footwear	0.61	0.56	0.05
	(0.02)	(0.02)	(0.03)
Mumias	0.56	0.57	-0.01
	(0.02)	(0.02)	(0.03)
Acres (owned and rented)	2.22	1.92	0.29
	(0.26)	(0.10)	(0.28)
Reads Swahili	0.91	0.91	0.00
	(0.01)	(0.01)	(0.02)
Had soil test	0.12	0.10	0.02
	(0.02)	(0.01)	(0.02)
Mentions Lime	0.03	0.05	-0.02
	(0.01)	(0.01)	(0.01)
Used Lime	0.06	0.07	-0.01
	(0.01)	(0.01)	(0.02)
Used fertilizer last LR season	0.84	0.84	0.00
	(0.02)	(0.02)	(0.03)
Grows legumes	0.81	0.83	-0.02
	(0.02)	(0.02)	(0.03)
Heard Lime	0.40	0.40	0.00
	(0.02)	(0.02)	(0.03)
Heard soil test	0.80	0.87	-0.07**
	(0.02)	(0.02)	(0.03)
Ever used DAP	0.94	0.94	0.00
	(0.01)	(0.01)	(0.02)
Ever used CAN	0.61	0.63	-0.02
	(0.02)	(0.02)	(0.03)
Ever used NPK	0.12	0.14	-0.02
	(0.02)	(0.02)	(0.02)
N	417	415	832
Joint F-Stat			1.06
P-value			0.386

Notes: The table shows summary statistics and balance tests using covariate variables from a baseline survey. Columns (1)–(2) display the mean and standard error of each characteristic for each treatment group. Column (3) displays the differences across columns and corresponding standard error. *Mumias* denotes share of farmers from Kakamega county (Mumias area), *Had soil test* denotes ever having a soil test, *Mentions Lime* is a dummy variable with value one if respondent mentioned lime as a strategy to reduce soil acidity. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table E2: IPA/PAD1-K: Summary Statistics & Balance

	Control (1)	General (2)	Specific (3)	(1) vs. (2) (4)	(1) vs. (3) (5)	(2) vs. (3) (6)
Age	46.25	46.01	45.59	0.25	0.66	0.42
	(0.49)	(0.45)	(0.43)	(0.66)	(0.65)	(0.63)
Female	0.37	0.37	0.37	-0.01	0.00	0.00
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Primary school	0.60	0.61	0.66	-0.01	-0.05*	-0.04
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Secondary school	0.10	0.10	0.10	0.00	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
MSC sample	0.53	0.53	0.53	0.00	0.00	0.00
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
pH prediction	5.42	5.40	5.40	0.02	0.02	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prefers English	0.30	0.27	0.30	0.03	0.00	-0.03
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Mentions lime	0.16	0.17	0.17	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Land size (acres)	2.00	1.86	2.14	0.14	-0.14	-0.28
	(0.09)	(0.08)	(0.31)	(0.12)	(0.32)	(0.32)
Has ever used lime	0.12	0.13	0.12	-0.01	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Used DAP last LR season	0.78	0.78	0.80	0.00	-0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Used NPK last LR season	0.04	0.05	0.04	-0.01	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Used CAN last LR season	0.62	0.62	0.59	0.00	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Used Fert.	0.18	0.18	0.18	0.00	0.00	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Used Mavuno last LR season	0.15	0.13	0.16	0.02	-0.01	-0.03
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Main network	0.95	0.94	0.94	0.01	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	632	633	632	1265	1264	1265
Joint F-Stat				0.57	0.78	0.75
P-value				0.909	0.714	0.746

Notes: The table shows summary statistics and balance tests using covariate variables from a baseline survey. Columns (1)–(3) display the mean and standard error of each characteristic for each treatment group. Columns (4)-(6) display the difference across columns and the corresponding standard error. *MSC Sample* denotes share of farmers from the Mumias Sugar Company sample. *pH prediction* represents the median pH level measured in the farmer's catchment area. *Mentions Lime* is a dummy variable with value one if respondent mentioned lime as a strategy to reduce soil acidity. Fertilizer use variables refer to input use during the 2016 long rains season. *Main network* indicates whether the farmer's phone service provider is the main network in area. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table E3: IPA/PAD2-K: Additional Summary Statistics & Balance

	Control (1)	SMS (2)	SMS+Call (3)	SMS+Call Offer (4)	(1) vs. (2) (5)	(1) vs. (3) (6)	(1) vs. (4) (7)
Age	42.10	41.40	41.48	41.44	0.70	0.61	0.66
	(0.32)	(0.31)	(0.32)	(0.31)	(0.45)	(0.46)	(0.45)
Female	0.34	0.34	0.34	0.34	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Primary school	0.72	0.70	0.69	0.71	0.01	0.02	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Secondary school	0.13	0.13	0.12	0.13	-0.01	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
pH prediction	5.37	5.37	5.37	5.37	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Prefers English	0.36	0.35	0.34	0.35	0.01	0.02	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Mentions lime	0.26	0.26	0.24	0.25	0.00	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Land size (acres)	2.02	1.85	2.09	2.03	0.17**	-0.07	-0.02
	(0.06)	(0.05)	(0.09)	(0.06)	(0.08)	(0.11)	(0.08)
Maize yield (t/ha)	1.51	1.46	1.37	1.49	0.05	0.15***	0.02
	(0.04)	(0.03)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)
Used Lime	0.09	0.09	0.09	0.10	-0.01	0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
OAF Participant	0.35	0.34	0.35	0.37	0.00	-0.01	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Used CAN last LR season	0.64	0.62	0.65	0.62	0.02	0.00	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Used Urea last LR season	0.18	0.20	0.20	0.18	-0.02	-0.02	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Used Mavuno last LR season	0.08	0.08	0.07	0.09	0.00	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Recommended Lime	0.77	0.76	0.77	0.76	0.01	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
N	1470	1475	1473	1472	2945	2943	2942
Joint F-Stat					0.93	0.64	0.51
P-value					0.52	0.84	0.93

Notes: The table shows summary statistics and balance tests using covariate variables from a baseline survey. Columns (1)–(4) display the mean and standard error of each characteristic for each treatment group. Columns (5)-(7) display the difference across columns and the corresponding standard error. *pH prediction* represents the median pH level measured in the farmer's ward used to provide lime recommendations. *OAF Participant* is dummy variable indicating whether the farmer has ever been enrolled in the OAF program. *Mentions Lime* is a dummy variable with value one if respondent mentioned lime as a strategy to reduce soil acidity. Fertilizer use variables refer to input use during the 2016 long rains season. *Recommended lime* indicates whether the farmer resided in a ward where lime was recommended. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table E4: OAF1-K: Additional Summary Statistics & Balance

	Control (1)	Broad (2)	Detailed (3)	(1) vs. (2) (4)	(1) vs. (3) (5)	(2) vs. (3) (6)
Female	0.64	0.64	0.67	0.00	-0.03	-0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Group size	9.08	9.24	9.07	-0.16	0.01	0.17*
	(0.07)	(0.07)	(0.07)	(0.10)	(0.10)	(0.10)
OAF Seasons	1.51	1.50	1.52	0.00	-0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Maize inputs (acres)	0.50	0.49	0.50	0.01	-0.01	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Repayment Incentive (hoe)	0.06	0.07	0.08	-0.01	-0.02***	-0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
pH prediction	5.48	5.48	5.48	0.00	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Intercropped (acres)	0.01	0.01	0.01	0.00*	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Extra CAN purchased	0.04	0.05	0.05	-0.01**	-0.02***	0.00
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Onions	0.09	0.13	0.12	-0.04***	-0.03**	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Storage Bags	0.23	0.31	0.24	-0.07*	-0.01	0.06
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Solar Lamp	0.44	0.45	0.46	-0.01	-0.02	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Health Insurance	0.22	0.23	0.21	-0.01	0.01	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	1559	1684	1641	3243	3200	3325
Joint F-Stat				1.74	1.93	1.71
P-value				0.053	0.026	0.058

Notes: The table shows summary statistics and balance tests using covariate variables from OAF long rain 2016 administrative records (before the trial took place). Columns (1)-(3) display mean and standard errors of each variable, by treatment group. Columns (4)-(6) display the difference across columns and the corresponding standard error. *Group size* denotes number of farmers in the participant's OAF group, *OAF seasons* denotes the number of seasons of enrollment in the OAF program, *Maize inputs (acres)* represents the size of maize inputs package purchased, *Repayment Incentive* is a dummy variable with value one if the farmer obtained a hoe as bonus for early repayment, *pH prediction* is the variable obtained using kriging interpolation that was used to produce detailed recommendations. *Intercropped* indicates the size of beans input package, for maize-beans intercropping, *Extra CAN*, *Onions*, *Solar Lamps*, and *Health Insurance* are dummy variables equal to one if the farmer purchased those additional products. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table E5: OAF2-K: Summary Statistics & Balance

	Control (1)	Lime only (2)	Lime + CAN (3)	(1) vs. (2) (4)	(1) vs. (3) (5)	(2) vs. (3) (6)
Age (years)	48.40	48.30	48.44	0.10	-0.04	-0.14
	(0.15)	(0.10)	(0.20)	(0.18)	(0.25)	(0.22)
Female	0.69	0.69	0.68	0.00	0.01	0.01
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Group size	9.87	9.82	9.92	0.04	-0.05	-0.10**
	(0.03)	(0.02)	(0.04)	(0.04)	(0.05)	(0.05)
OAF Seasons	2.23	2.23	2.22	0.00	0.01	0.01
	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)
Maize inputs (acres)	0.51	0.51	0.53	0.00	-0.01**	-0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
pH prediction	5.33	5.33	5.33	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Intercropped (acres)	0.01	0.01	0.01	0.00*	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Extra CAN purchased	0.15	0.15	0.14	0.00	0.01	0.01
	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)
Onions	0.04	0.04	0.04	0.00	-0.01	-0.01*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Storage Bags	0.40	0.42	0.40	-0.02	0.00	0.02
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Solar Lamp	0.42	0.42	0.43	0.00	0.00	-0.01
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Credit size	9504.07	9467.78	9616.95	36.29	-112.88	-149.17**
	(49.84)	(31.67)	(64.63)	(58.68)	(81.55)	(71.17)
N	8142	19558	4872	27700	13014	24430
Joint F-Stat				0.63	0.86	1.78
P-value				0.816	0.583	0.045

Notes: The table shows summary statistics and balance tests using covariate variables from OAF long rain 2017 administrative records (before the trial took place). Columns (1)-(3) display mean and standard errors of each variable, by treatment group. Columns (4)-(6) display the difference across columns and the corresponding standard error. *Group size* denotes number of farmers in the participant's OAF group, *OAF seasons* denotes the number of seasons of enrollment in the OAF program, *Maize inputs (acres)* represents the size of maize inputs package purchased, *pH prediction* was obtained using kriging interpolation. *Intercropped* indicates the size of beans input package, for maize-beans intercropping, *Extra CAN*, *Onions*, *Solar Lamps*, are dummy variables equal to one if the farmer purchased those additional products. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table E6: OAF3-R: Summary Statistics & Balance

	G0: No SMS	G1: Same SMS	G2: Diff SMS	G3-Control	G3-Treated	(1) vs. (2)	(1) vs. (3)	(1) vs. (4)	(1) vs. (5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Group size	10.73	10.70	10.81	10.74	10.73	0.03	-0.08	-0.01	0.00
	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.09)	(0.09)	(0.07)	(0.07)
OAF Seasons	2.01	2.02	2.01	2.02	2.01	-0.01	0.01	-0.00	-0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Bought lime	0.06	0.07	0.06	0.06	0.06	-0.00	0.00	0.01^{*}	0.01^{*}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Planting fertilizer (kg)	13.78	13.87	13.80	13.74	13.80	-0.09	-0.02	0.04	-0.02
	(0.20)	(0.20)	(0.19)	(0.13)	(0.14)	(0.28)	(0.28)	(0.24)	(0.24)
Bought urea	0.75	0.75	0.75	0.74	0.75	-0.01	0.00	0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Solar Lamp	0.28	0.26	0.27	0.28	0.28	0.01^{*}	0.00	-0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Credit size	22768.25	22390.75	22739.35	22983.56	22941.68	377.51	28.90	-215.31	-173.43
	(213.22)	(215.49)	(222.56)	(157.12)	(157.15)	(303.13)	(308.19)	(264.84)	(264.86)
N	19743	18988	18821	28520	28497	38731	38564	48263	48240
Joint F-Stat						0.62	0.30	0.81	0.55
P-value						0.74	0.96	0.58	0.80

Notes: The table shows summary statistics and balance tests using covariate variables from OAF 2017 administrative records (before the trial took place). Columns (1) - (5) display mean and standard errors of each variable, by treatment group. Columns (6)-(9) displays the difference across columns and the corresponding standard error. *Group size* denotes number of farmers in the participant's OAF group, *OAF seasons* denotes the number of seasons of enrollment in the OAF program, *Bought lime* is a dummy indicating whether the farmer purchased lime. *Planting fertilizer* indicates the quantity of planting fertilizer (DAP and NPK) purchased, and *Bought urea* is a dummy indicating whether the farmer purchased urea. *Solar Lamps* is a dummy variables equal to one if the farmer purchased any solar lamps. *Credit size* reports the size of the OAF loan. Standard errors are clustered at the farmer group level. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table E7: Probability of Collecting Data

	LPM			Odd ratios		
	Survey	Enroll 1st	Enroll 2nd	Survey	Enroll 1st	Enroll 2nd
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. KALRO						
Treated	0.019			1.325		
	(0.018)			(0.358)		
Mean Control	0.919			0.919		
Observations	833			833		
Panel B. IPA/PAD1-K						
Treated	0.014			1.086		
	(0.020)			(0.126)		
Mean Control	0.79			0.79		
Observations	1897			1897		
Panel C. IPA/PAD2-K						
Treated	-0.002			0.985		
	(0.012)			(0.077)		
Mean Control	0.82			0.82		
Observations	5890			5890		
Panel D. OAF1-K						
Treated	-0.018	-0.002	0.014	0.904	0.991	1.060
	(0.013)	(0.015)	(0.015)	(0.066)	(0.062)	(0.066)
Mean Control	0.25	0.60	0.40	0.25	0.60	0.40
Observations	4884	4884	4884	4884	4884	4884
Panel E. OAF2-K						
Treated		0.002	0.007		1.009	1.029
		(0.005)	(0.006)		(0.030)	(0.027)
Mean Control		0.76	0.56		0.76	0.56
Observations		32572	32572		32572	32572
Panel F. OAF3-R						
Treated		0.010	0.004		1.043	1.016
		(0.007)	(0.007)		(0.031)	(0.030)
Mean Control		0.65	0.48		0.65	0.48
Observations		86049	86049		86049	86049

Notes: The dependent variable in Panel A takes the value of one if the farmer completed the in person endline survey. In panels B and C the dependent variable indicates whether the farmer completed the phone-based endline survey. In panel D the dependent variable in columns (1) and (4) is a dummy variable indicating whether the farmer completed a phone-based survey (conducted with 30% of original sample). In panels D-F, columns (2) and (5) the dependent variable indicates whether the farmer enrolled in the OAF program (i.e. placed an input order) in the season in which the program took place, while in columns (3) and (6) the dependent variable indicates whether they enrolled in the program in the following year. Columns (1)-(3) report marginal effects estimated using OLS, columns (4)-(6) report odds ratios estimated using Logit. * $p < .10$, ** $p < .05$, *** $p < .01$.

F Results by Experiment: Pooled Treatment Arms

Table F1: Awareness and Knowledge about Lime

	LPM		Odds ratios	
	Heard Lime	Knows Lime Use	Heard Lime	Knows Lime Use
	(1)	(2)	(3)	(4)
Panel A. KALRO				
Treated	-0.004	0.023	0.968	1.151
	(0.032)	(0.024)	(0.170)	(0.279)
Mean Control	0.58	0.14	0.58	0.14
Observations	773	773	773	773
Panel B. IPA/PAD1-K				
Treated	0.034	0.086***	1.257	1.598***
	(0.022)	(0.025)	(0.191)	(0.209)
Mean Control	0.78	0.33	0.77	0.33
Observations	1471	1471	1435	1471
Panel C. IPA/PAD2-K				
Treated	0.054***	0.099***	1.601***	1.591***
	(0.012)	(0.016)	(0.153)	(0.117)
Mean Control	0.81	0.45	0.81	0.45
Observations	4822	4822	4638	4771
Panel D. OAF1-K				
Treated	0.001	0.100***	0.998	1.629***
	(0.025)	(0.030)	(0.174)	(0.237)
Mean Control	0.80	0.32	0.80	0.32
Observations	1087	1087	1087	1087

Notes: *Heard Lime* is a dummy variable reporting whether farmers had heard about agricultural lime before. *Knows Lime Use* is coded as one if farmer mentions lime a strategy to deal with or reduce soil acidity. All regressions include controls. Columns (1) and (2) report marginal effects estimated using OLS, columns (3) and (4) report odds rations estimated using Logit. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table F2: Followed Lime Recommendations

	LPM			Logit		
	Survey	Admin (all)	Admin (enrol)	Survey	Admin (all)	Admin (enrol)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. KALRO						
Treated	-0.003			0.963		
	(0.020)			(0.281)		
Mean Control	0.10			0.14		
Observations	773			561		
Panel B. IPA/PAD1-K						
Treated	0.041**	0.019		1.617**	1.164	
	(0.017)	(0.017)		(0.322)	(0.173)	
Mean Control	0.22	0.24		0.22	0.25	
Observations	1471	1897		1378	1854	
Panel C. IPA/PAD2-K						
Treated	0.075***	0.030***		1.653***	1.379***	
	(0.013)	(0.009)		(0.150)	(0.145)	
Mean Control	0.31	0.30		0.31	0.28	
Observations	4822	5890		4641	5476	
Panel D. OAF1-K						
Treated	0.041*	0.030***	0.054***	1.564**	1.395***	1.495***
	(0.021)	(0.009)	(0.014)	(0.328)	(0.145)	(0.164)
Mean Control	0.12	0.10	0.17	0.12	0.10	0.17
Observations	1087	4884	2931	1087	4884	2931
Panel E. OAF2-K						
Treated		0.025***	0.031***		1.153***	1.201***
		(0.005)	(0.006)		(0.035)	(0.043)
Mean Control		0.32	0.42		0.32	0.42
Observations		32572	24825		32572	24623
Panel F. OAF3-R						
Treated		0.008***	0.012***		1.216***	1.227***
		(0.002)	(0.003)		(0.059)	(0.060)
Mean Control		0.05	0.07		0.07	0.10
Observations		86049	56303		59896	40958

Notes: This table reports the marginal effect of each program on whether farmers followed the lime recommendations. Columns (1)-(3) report marginal effects estimated using OLS. Columns (4)-(6) report odds ratios, estimated using Logit. Columns (1) and (4) report survey result. Column (2) and (5) shows results for the administrative data (lime purchases or coupon redemption) for the entire sample of farmers participating int the experiment. Columns (3) and (6) show results for the administrative data for the subset of OAF farmers registered in the program in that season. In panels A and D-F the dependent variable takes value one if the farmer used or acquired agricultural lime. In panels B and C, the dependent variable takes the value one if farmer used lime in an area where it was recommended, or did not use lime in an area where it was not recommended. All regressions include controls. Robust standard errors shown in parenthesis. In panel F standard errors are clustered at the OAF group level. $^{*}p < .10$, ** $p < .05$, *** $p < .01$.

Table F3: Persistence & Re-treatment

	Followed lime recommendations 2nd season					
	LPM			Logit		
	Survey	Admin (all)	Admin (enrol)	Survey	Admin (all)	Admin (enrol)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. KALRO						
Treated (S1)		-0.006			0.904	
		(0.022)			(0.238)	
Mean Control		0.11			0.13	
Observations		773			664	
Panel B. IPA/PAD1-K						
Treated (S1 & S2)	0.051***	0.006		1.570***	1.121	
	(0.019)	(0.010)		(0.273)	(0.273)	
Mean Control	0.15	0.11		0.16	0.07	
Observations	1471	1897		1409	1531	
Panel C. IPA/PAD2-K						
Treated (S1)	0.010*			1.374		
	(0.006)			(0.267)		
Mean Control	0.22			0.19		
Observations	2566			1745		
Panel D. OAF1-K						
Treated (S1)		0.004	-0.011		1.088	0.912
		(0.014)	(0.021)		(0.227)	(0.194)
Treated (S2)		0.028	0.032		1.472*	1.388
		(0.018)	(0.028)		(0.345)	(0.333)
Treated (S1 & S2)		0.023	0.018		1.405*	1.229
		(0.015)	(0.022)		(0.289)	(0.257)
Mean Control		0.08	0.12		0.08	0.12
Observations		2931	1986		2931	1986
Panel E. OAF2-K						
Treated (S1 & S2)		0.002	0.000		1.025	1.001
		(0.004)	(0.005)		(0.058)	(0.060)
Mean Treated S2 only		0.09	0.12		0.09	0.13
Observations		24825	18356		24380	17997
Panel F. OAF3-R						
Treated (S1)		0.006*	0.008*		1.098*	1.094
		(0.004)	(0.005)		(0.062)	(0.063)
Treated (S2)		0.020***	0.026***		1.345***	1.326***
		(0.005)	(0.007)		(0.100)	(0.102)
Treated (S1 & S2)		0.017***	0.023***		1.308***	1.288***
		(0.004)	(0.005)		(0.078)	(0.078)
Mean Control		0.08	0.11		0.09	0.12
Observations		51923	36012		41362	32051

Notes: This table reports the effect of each program on whether farmers followed the lime recommendations for a second season. *Treated (S1)* indicates that the farmer received SMS only in the first season, *Treated (S2)* indicates that the farmer received SMS only in the second season, *Treated (S1 & S2)* indicates that the farmer received SMS in both season. In panel E the comparison group is the set of farmers that received messages only in the second season. Columns (1)-(3) report marginal effects estimated using OLS. Columns (4)-(6) report odds ratios, estimated using Logit. Columns (1) and (4) report survey result. Column (2) and (5) shows results for the administrative data (lime purchases or coupon redemption) for the entire sample of farmers participating int the experiment. Columns (3) and (6) show results for the administrative data for the subset of OAF farmers registered in the program in the second season. In panels D, E, and F, the sample is restricted to the farmers registered for the program in the first season as the others were not eligible for receiving SMS messages in the second season. In panels A and D-F the dependent variable takes value one if the farmer used or acquired agricultural lime. In panels B and C, the dependent variable takes the value one if farmer used lime in an area where it was recommended, or did not use lime in an area where it was not recommended. All regressions include controls. Robust standard errors shown in parenthesis. In panel F standard errors are clustered at the OAF group level. $^{*}p < .10$, ** $p < .05$, *** $p < .01$.

Table F4: Use of Recommended Fertilizers

	LPM		Odds Ratios	
	Survey	Admin	Survey	Admin
	(1)	(2)	(3)	(4)
Panel A. KALRO				
Treated	-0.054*		0.747*	
	(0.032)		(0.128)	
Mean Control	0.50		0.52	
Observations	773		756	
Panel B. IPA/PAD1-K				
Treated	0.005	0.011	1.063	1.695
	(0.020)	(0.007)	(0.177)	(0.617)
Mean Control	0.15	0.02	0.17	0.03
Observations	1471	1897	1371	1278
Panel C. IPA/PAD2-K				
Treated	0.034***	0.005	1.298***	1.244
	(0.013)	(0.005)	(0.124)	(0.256)
Mean Control	0.16	0.02	0.16	0.04
Observations	4822	5890	4669	3471
Panel D. OAF2-K				
Treated		0.023***		1.255***
		(0.005)		(0.061)
Mean Control		0.14		0.14
Observations		32572		32572

Notes: This table reports the effect of each program on use chemical fertilizers. In panel A, the dependent variable takes value one if the farmer used recommended topdressing fertilizer (CAN or Mavuno). In panel B and C, the dependent variable in columns (1) and (3) indicates whether the farmer used urea, while the dependent variable in columns (2) and (4) indicate whether the used the electronic coupon to purchase urea. In panel D, the dependent variable indicates whether the farmer purchased “Extra CAN” from OAF. Since only a subset of treated farmers were recommended Extra CAN, here *Treated* indicates that the farmer was assigned to the “Lime+CAN” subtreatment. The outcomes reported in odd columns are measured using using survey data, while the outcomes reported in even columns are measured administrative data from coupon redemption or purchases from OAF. All regressions include controls. Robust standard errors in parentheses. Columns (1) and (2) report marginal effects measured using OLS, columns (3) and (4) report odds ratios measured using Logit. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table F5: Use of All Recommended Inputs and Other Inputs

	Recommended Inputs (index) (1)	Other Inputs (Index) (2)
Panel A. KALRO		
Treated	-0.027	0.019
	(0.033)	(0.039)
Observations	773	773
Panel B. IPA/PAD1-K		
Treated	0.058*	-0.018
	(0.033)	(0.030)
Observations	1471	1471
Panel C. IPA/PAD2-K		
Treated	0.057***	-0.011
	(0.017)	(0.017)
Observations	4822	4822
Panel D. OAF1-K		
Treated	0.089***	0.020
	(0.027)	(0.013)
Observations	4884	4884
Panel E. OAF2-K		
Treated	0.046***	0.007
	(0.009)	(0.006)
Observations	32572	32572
Panel F. OAF3-R		
Treated	0.036***	0.003
	(0.009)	(0.007)
Observations	86049	86049

Notes: This table present results of indexes of recommended inputs (column (1)) and other inputs not mentioned by the SMS messages (column (2)). Each index is composed of different variables depending on the project. For a full list of variables see table D1. The coefficients are average effect sizes. All regressions include controls (panel F includes fixed effect at the OAF sector level instead of the site level). Robust standard errors in parentheses. In panel F standard errors are clustered at the OAF group level. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table F6: Amount of Agricultural Lime by Type of Recommendation

	Kg Lime (admin)	
	Lime Rec	Lime not Rec
	(1)	(2)
Panel A. IPA/PAD1-K		
Treated	0.119	1.379
	(0.618)	(1.233)
Mean Control	2.85	3.32
Observations	1552	345
Panel B. IPA/PAD2-K		
Treated	0.966**	-1.495*
	(0.444)	(0.768)
Mean Control	3.52	3.56
Observations	4512	1378
Panel C. OAF1-K		
Treated	3.207***	
	(0.794)	
Mean Control	5.82	
Observations	4884	
Panel D. OAF2-K		
Treated	2.237***	
	(0.449)	
Mean Control	17.90	
Observations	32572	
Panel E. OAF3-R		
Treated	0.296*	
	(0.169)	
Mean Control	2.55	
Observations	86049	

Notes: The table reports the effects of the programs on quantity of lime purchased, expressed in kgs. All regressions include controls. Robust standard errors in parentheses. In panel E the standard errors are clustered at the OAF group level. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table F7: Spillovers

	LPM				Odds Ratios			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. OAF1-K								
N treated	-0.001				0.985			
	(0.006)				(0.070)			
Sample	Control				Control			
Observations	1559				1453			
Panel B. OAF2-K								
N treated	0.006				1.038*			
	(0.004)				(0.024)			
Sample	Control				Control			
Observations	8142				7966			
Panel C. OAF3-R								
N treated	-0.000		0.000		1.004		1.014	
	(0.001)		(0.000)		(0.017)		(0.011)	
Group Treated		0.004***		0.002		1.130**		1.101
		(0.001)		(0.001)		(0.054)		(0.077)
Control Mean		0.03	0.02	0.02		0.04	0.04	0.04
Sample	G3-Control	G3-C& G0	All	All	G3-Control	G3-C & G0	All	All
Has Phone	Yes	Yes	No	No	Yes	Yes	No	No
Observations	28520	90788	101891	101891	22093	72819	61448	61448

Notes: This table reports spillovers effects at the OAF group level. The dependent variable indicates whether the farmer has purchased lime. *N treated* indicates the number of treated farmers in the OAF group, *Group treated* is dummy equal to 1 if some farmers in the group were treated. The sample is restricted to farmers that were not assigned to receive messages or could not receive them because they did not have a valid phone number registered in the OAF database. All regressions include controls. Robust standard errors in parentheses, in panel C, columns (2)-(4) and (6)-(8) standard errors are clustered at the OAF group level. Columns (1)-(4) report marginal effects measured using OLS, columns (5)-(8) report odds ratios measured using Logit. * $p < .10$, ** $p < .05$, *** $p < .01$.

G Results by Experiment: By Treatment Arms

Table G1: Knowledge and Adoption by Treatment Arms

	LPM				Odds ratios			
	Heard Lime	Knows Lime	Followed Lime Rec		Heard Lime	Knows Lime	Followed Lime Rec	
			Survey	Admin			Survey	Admin
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. IPA/PAD1-K								
General	0.037	0.061**	0.032	0.020	1.282	1.416**	1.455	1.188
	(0.025)	(0.029)	(0.020)	(0.020)	(0.226)	(0.213)	(0.337)	(0.206)
Specific	0.030	0.111***	0.050**	0.017	1.232	1.793***	1.785***	1.141
	(0.025)	(0.030)	(0.020)	(0.019)	(0.218)	(0.268)	(0.391)	(0.190)
Mean Control	0.78	0.33	0.22	0.24	0.77	0.33	0.22	0.25
Observations	1471	1471	1471	1897	1435	1471	1378	1854
p-value General=Specific	0.751	0.095	0.389	0.863	0.826	0.105	0.330	0.807
Panel B. IPA/PAD2-K								
SMS	0.042***	0.092***	0.067***	0.030***	1.414***	1.533***	1.569***	1.390***
	(0.015)	(0.020)	(0.016)	(0.012)	(0.168)	(0.137)	(0.168)	(0.172)
SMS + Call	0.070***	0.116***	0.095***	0.020*	1.881***	1.730***	1.870***	1.247*
	(0.014)	(0.019)	(0.017)	(0.012)	(0.231)	(0.156)	(0.205)	(0.162)
SMS + Call Offer	0.051***	0.089***	0.064***	0.039***	1.568***	1.520***	1.539***	1.507***
	(0.015)	(0.020)	(0.016)	(0.012)	(0.192)	(0.139)	(0.168)	(0.187)
Mean Control	0.81	0.45	0.31	0.30	0.81	0.45	0.31	0.28
Observations	4822	4822	4822	5890	4638	4771	4641	5476
p-value SMS=SMS+Call	0.041	0.206	0.109	0.369	0.026	0.178	0.092	0.373
p-value SMS=SMS+Call Offer	0.553	0.882	0.870	0.474	0.416	0.927	0.854	0.484
p-value SMS+Call=SMS+Call Offer	0.155	0.162	0.077	0.107	0.167	0.159	0.065	0.117
Panel C. OAF1-K								
Broad	0.006	0.091**	0.038	0.026**	1.019	1.573***	1.530*	1.342**
	(0.029)	(0.036)	(0.025)	(0.011)	(0.206)	(0.264)	(0.370)	(0.159)
Detailed	-0.003	0.109***	0.045*	0.034***	0.977	1.689***	1.598**	1.450***
	(0.030)	(0.035)	(0.025)	(0.011)	(0.199)	(0.279)	(0.372)	(0.171)
Mean Control	0.80	0.32	0.12	0.10	0.80	0.32	0.12	0.10
Observations	1087	1087	1087	4884	1087	1087	1087	4884
p-value Broad=Detailed	0.764	0.631	0.799	0.501	0.836	0.661	0.846	0.485
Panel D. OAF2-K								
Lime only				0.023***				1.141***
				(0.005)				(0.036)
Lime+CAN				0.032***				1.202***
				(0.008)				(0.052)
Mean Control				0.32				0.32
Observations				32572				32572
p-value Lime only=Lime+CAN				0.174				0.178
Panel E. OAF3-R								
G1: Same SMS				0.006**				1.160**
				(0.002)				(0.070)
G2: Diff SMS				0.009***				1.240***
				(0.003)				(0.076)
G3-Treated				0.009***				1.237***
				(0.002)				(0.066)
Mean Control				0.05				0.07
Observations				86049				59896
p-value G1=G2				0.257				0.262

Notes: The table shows the effect of each of the main treatments on knowledge of lime and probability to follow recommendations. The dependent variable in column (1) is a dummy variable reporting whether farmers had heard about agricultural lime before. The dependent variable in column (2) is coded as one if farmer mentions lime a strategy to deal with or reduce soil acidity. The dependent variable in columns (3) and (4) indicates whether farmers followed lime recommendations. In panels A and B, it takes value one if the farmer used lime and lime was recommended or if farmer did not use lime and lime was not recommended, zero otherwise. In panels C-E takes value one if the farmer used lime, zero otherwise. All regressions include controls. Robust standard errors in parenthesis. In panel E the standard errors are clustered at the OAF group level. Columns (1) - (4) report marginal effects estimated using OLS, columns (5) - (8) report odds ratios estimated using Logit. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table G2: Number of Messages

	Purchased lime					
		LPM			Odds ratios	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. OAF2-K						
N Lime SMS	0.006***			1.035***		
	(0.001)			(0.008)		
N Lime SMS$\geq$ 1		-0.003			0.983	
		(0.012)			(0.068)	
N Lime SMS $\geq$ 2		0.025**	0.025**		1.159**	1.157**
		(0.012)	(0.013)		(0.083)	(0.083)
N Lime SMS $\geq$ 3		0.004	0.004		1.023	1.023
		(0.008)	(0.008)		(0.046)	(0.045)
N Lime SMS $\geq$ 4		0.004	0.004		1.023	1.022
		(0.008)	(0.008)		(0.045)	(0.045)
N Lime SMS = 5		-0.005	-0.005		0.973	0.974
		(0.008)	(0.008)		(0.044)	(0.043)
Mean Control	0.32	0.32		0.32	0.32	
Observations	32572	32572	24430	32572	32572	24430
Includes Control Group	Yes	Yes	No	Yes	Yes	No
Panel B. OAF3-R						
N Lime SMS	0.003***			1.066***		
	(0.001)			(0.013)		
N Lime SMS$\geq$ 1		0.003			1.077	
		(0.002)			(0.064)	
N Lime SMS $\geq$ 2		0.007***	0.007***		1.173***	1.162***
		(0.002)	(0.002)		(0.063)	(0.061)
N Lime SMS $\geq$ 3		-0.001	-0.001		0.981	0.991
		(0.002)	(0.002)		(0.051)	(0.050)
N Lime SMS = 4		0.002	0.002		1.038	1.026
		(0.002)	(0.002)		(0.055)	(0.053)
Mean Control	0.05	0.05		0.07	0.07	
Observations	86049	86049	66306	59896	59896	45935
Includes Control Group	Yes	Yes	No	Yes	Yes	No

Notes: The table shows the effect number of messages on lime purchases. The dependent variable in column indicates whether farmers purchased lime from OAF. All regressions include controls (panel B column (6) includes fixed effect at the OAF sector level instead of the site level). Robust standard errors in parenthesis. In panel B the standard errors are clustered at the OAF group level. Columns (1) - (4) report marginal effects estimated using OLS, columns (5) - (8) report odds ratios estimated using Logit. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table G3: Message Framing

	Purchased lime					
		LPM			Odds ratios	
	(1)	(2)	(3)	(5)	(6)	(7)
Panel A. OAF2-K						
Basic	0.017**			1.108**		
	(0.008)			(0.051)		
Yield Increase	0.034***	0.017*		1.217***	1.097*	
	(0.008)	(0.009)		(0.056)	(0.057)	
Experimentation (self)	0.027***	0.010		1.171***	1.056	
	(0.008)	(0.009)		(0.054)	(0.055)	
Experimentation (neighbors)	0.013*	-0.004		1.079	0.973	
	(0.008)	(0.009)		(0.050)	(0.051)	
Social Compasison	0.028***	0.010		1.174***	1.058	
	(0.008)	(0.009)		(0.054)	(0.056)	
Self-efficacy	0.028***	0.010		1.175***	1.059	
	(0.008)	(0.009)		(0.054)	(0.056)	
Family framed SMS			-0.016***			0.912***
			(0.005)			(0.028)
Mean Control	0.32			0.32		
Observations	32572	24430	24430	32572	24430	24430
Includes Control Group	Yes	No	No	Yes	No	No
Panel B. OAF3-R						
General promotion	0.011***			1.300***		
	(0.003)			(0.088)		
Specific + yield impact	0.009***	-0.002		1.240***	0.964	
	(0.003)	(0.003)		(0.084)	(0.070)	
Self-diagnosis	0.010***	-0.001		1.267***	1.002	
	(0.003)	(0.003)		(0.086)	(0.073)	
Soil test	0.007***	-0.003		1.194**	0.936	
	(0.003)	(0.003)		(0.083)	(0.068)	
How travertine works	0.006**	-0.005		1.157**	0.900	
	(0.003)	(0.003)		(0.080)	(0.066)	
Order immediately	0.006**	-0.005		1.177**	0.919	
	(0.003)	(0.003)		(0.083)	(0.068)	
Your cell is acidic + yield impact	0.007**	-0.004		1.180**	0.942	
	(0.003)	(0.003)		(0.081)	(0.069)	
SMS framed as gain			0.000			1.029
			(0.002)			(0.041)
Mean Control	0.05			0.07		
Observations	86049	66306	66306	59896	45935	45935
Includes Control Group	Yes	No	No	Yes	No	No

Notes: The table shows the effect different messages framing of messages on lime purchases. The dependent variable in column indicates whether farmers purchased lime from OAF. All regressions include controls (panel B columns (6) and (7) include fixed effect at the OAF sector level instead of the site level). Robust standard errors in parenthesis. In panel B the standard errors are clustered at the OAF group level. * $p < .10$, ** $p < .05$, *** $p < .01$.

H Heterogeneity by Experiment

Table H1: Heterogeneity in Following Lime Recommendations

	Odds ratios					
[X]	Female	Primary School	Large Farm	Young	Used Input	Heard Input
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. KALRO						
Treated	0.446*	1.114	1.121	1.446	0.955	0.678
	(0.212)	(0.487)	(0.481)	(0.635)	(0.346)	(0.385)
[X]	0.537	1.494	1.537	1.240	26.883***	3.912***
	(0.231)	(0.628)	(0.856)	(0.541)	(18.258)	(1.730)
[X] *Treated	3.336**	0.748	0.704	0.445	0.538	1.604
	(1.973)	(0.432)	(0.448)	(0.276)	(0.423)	(1.041)
Mean Control	0.14	0.14	0.14	0.14	0.14	0.14
Observations	561	561	561	561	561	561
Panel B. IPA/PAD1-K						
Treated	1.222	0.993	1.202	1.343	1.062	1.179
	(0.226)	(0.249)	(0.241)	(0.263)	(0.165)	(0.190)
[X]	0.872	0.859	1.017	1.153	0.806	1.061
	(0.234)	(0.231)	(0.290)	(0.366)	(0.332)	(0.350)
[X] *Treated	0.870	1.277	0.928	0.684	1.642	0.914
	(0.270)	(0.403)	(0.283)	(0.208)	(0.735)	(0.377)
Mean Control	0.25	0.25	0.25	0.25	0.25	0.25
Observations	1854	1854	1854	1854	1854	1854
Panel C. IPA/PAD2-K						
Treated	1.445***	1.355*	1.221*	1.292**	1.322***	1.226**
	(0.158)	(0.237)	(0.141)	(0.165)	(0.123)	(0.125)
[X]	1.316*	1.114	0.783	1.051	0.992	0.878
	(0.215)	(0.206)	(0.131)	(0.201)	(0.288)	(0.164)
[X] *Treated	0.768	0.966	1.236	1.046	0.982	1.344
	(0.143)	(0.197)	(0.225)	(0.186)	(0.319)	(0.280)
Mean Control	0.30	0.30	0.30	0.30	0.30	0.30
Observations	5732	5732	5732	5732	5732	5732
Panel D. OAF1-K						
Treated	1.520**	1.478	1.391***	1.577		
	(0.271)	(0.440)	(0.164)	(0.443)		
[X]	1.070	0.745	0.766	0.725		
	(0.201)	(0.263)	(0.195)	(0.266)		
[X] *Treated	0.880	1.120	1.013	0.988		
	(0.193)	(0.459)	(0.253)	(0.425)		
Mean Control	0.10	0.11	0.10	0.11		
Observations	4812	1151	4884	1151		
Panel E. OAF2-K						
Treated	1.142***		1.135***	1.112***		
	(0.059)		(0.036)	(0.037)		
[X]	1.302***		0.797***	0.760***		
	(0.070)		(0.056)	(0.041)		
[X]*Treated	0.977		0.952	1.033		
	(0.060)		(0.062)	(0.064)		
Mean Control	0.33		0.32	0.33		
Observations	31597		32572	31604		
Panel F. OAF3-R						
Treated			1.150**		1.211***	
			(0.066)		(0.062)	
[X]			1.290***		2.327***	
			(0.104)		(0.258)	
[X] *Treated			1.136		1.023	
			(0.096)		(0.127)	
Mean Control			0.07		0.07	
Observations			59896		59896	

Notes: This table shows results of heterogeneity analysis by sample. The dependent variable is whether the farmer followed the lime recommendations in the first season. Panel A reports survey results, panel B-F report results for the administrative data. We show results for gender, whether respondent completed primary school, whether the respondent's land is large (defined as above median use of inputs for the OAF samples and more than 1.5 acres of land for the other programs), whether the respondent was under 40 years old, whether the respondent had previously used the input, and whether the respondent had previous knowledge of the input. All regressions include controls (panel E does not include site FEs). Effect sizes are reported in terms of odds ratios measured using Logit. Robust standard errors in parentheses. In panel F standard errors are clustered at the OAF group level. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table H2: Heterogeneity in Following Fertilizer Recommendations

	Odds ratios					
$[X]$	Female	Primary School	Large Farm	Young	Used Input	Heard Input
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. KALRO						
Treated	0.782	0.734	0.755	0.878	0.531**	1.464
	(0.227)	(0.183)	(0.170)	(0.229)	(0.164)	(1.049)
$[X]$	0.818	1.171	1.430	1.489	2.174***	4.376**
	(0.218)	(0.295)	(0.497)	(0.391)	(0.597)	(2.546)
$[X]$ *Treated	0.931	1.018	0.969	0.730	1.566	0.490
	(0.340)	(0.349)	(0.340)	(0.256)	(0.598)	(0.363)
Mean Control	0.52	0.52	0.52	0.52	0.52	0.52
Observations	756	756	756	756	756	756
Panel B. IPA/PAD1-K						
Treated	2.124	1.394	2.036	1.629	1.548	
	(0.992)	(0.843)	(1.001)	(0.702)	(0.650)	
$[X]$	1.345	1.042	1.690	0.494	2.026	
	(0.905)	(0.710)	(1.120)	(0.425)	(1.519)	
$[X]$ *Treated	0.535	1.338	0.632	1.209	1.402	
	(0.410)	(1.019)	(0.473)	(1.016)	(1.180)	
Mean Control	0.03	0.03	0.03	0.03	0.03	
Observations	1278	1278	1278	1278	1278	
Panel C. IPA/PAD2-K						
Treated	1.011	1.512	1.443	1.145	1.454	
	(0.237)	(0.644)	(0.390)	(0.316)	(0.341)	
$[X]$	0.695	1.641	1.389	0.919	2.796***	
	(0.275)	(0.715)	(0.507)	(0.391)	(1.095)	
$[X]$ *Treated	1.737	0.755	0.655	1.103	0.487	
	(0.777)	(0.364)	(0.260)	(0.442)	(0.216)	
Mean Control	0.03	0.03	0.03	0.03	0.03	
Observations	4024	4024	4024	4024	4024	
Panel D. OAF2-K						
Treated	1.060		1.131***	1.153***	1.108**	
	(0.081)		(0.052)	(0.055)	(0.056)	
$[X]$	1.102		0.781**	0.867*	8.706***	
	(0.087)		(0.080)	(0.070)	(0.637)	
$[X]$*Treated	1.122		1.064	0.997	1.103	
	(0.102)		(0.103)	(0.092)	(0.093)	
Mean Control	0.15		0.14	0.15	0.14	
Observations	31597		32572	31604	32572	

Notes: This table shows results of heterogeneity analysis by sample. The dependent variable is whether the farmer followed the fertilizer recommendations in the first season. Panel A reports survey results, panel B-D report results for the administrative data. We show results for gender, whether respondent completed primary school, whether the respondent's land is large (defined as above median use of inputs for the OAF samples and more than 1.5 acres of land for the other programs), whether the respondent was under 40 years old, whether the respondent had previously used the input, and whether the respondent had previous knowledge of the input. All regressions include controls (panel D does not not include site FEs). Effect sizes are reported in terms of odds ratios measured using Logit. Robust standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

I Additional Notes on Metanalysis

- Cochran's Q statistic. The Q statistic is the weighed sum of squared differences between the individual studies and the mean effect across studies, where the weights are the same as in equation 4. In other words:

$$Q = \sum_{j=1}^{s} w_j (T_j - \frac{\sum_{j=1}^{s} w_j T_j}{\sum_{j=1}^{s} w_j})^2$$

  The Q statistic is a chi-square statistic with s (number of studies) minus 1 degrees of freedom. The null is that all treatments are equally effective. One concern with the Q statistic is that it has low power, particularly when the number of studies is small.

- Higgin's and Thompson's I^2. The percentage of variability, I^2, measures the share of variability not explained by sampling error and is given by:

$$I^2 = max\left\{0, \frac{Q-(s-1)}{Q}\right\}$$

  I^2 is less sensitive to the number of studies included, but it depends on their precision (Borenstein et al., 2017). While there is subjectivity on interpreting the magnitudes, Higgins et al. (2003) provides the following rules of thumb: I^2=25% for low, I^2=50% for moderate, and I^2=75% for high heterogeneity.We report I^2 and a corresponding 95% confidence interval.

- We also explore the role of covariates in explaining the between-study heterogeneity, by conducting random-effects meta-regressions:

$$\hat{\mu} \sim N(a^m + \beta^m x_j, \tau^2 + \hat{\sigma}^2)$$

  where β^m is the effect of covariate x on the mean impact of the intervention. We complement this analysis running heterogeneity specifications pooling all different datasets together.