The impact of digital agricultural extension service: Experimental evidence from rice farmers in India*

Shawn Cole
Harvard Business School

Jessica Goldberg
University of Maryland

Tomoko Harigaya
Precision Development

Jessica Zhu
Precision Development

May 27, 2025

Abstract

We evaluate at scale the impact of a digital agricultural advisory service reaching millions of smallholder farmers, in an eastern state of India. We randomized the rollout of the service among 13,675 rice farmers within five districts, and measured the impact on agricultural outcomes using both survey and remote sensing data. Using survey data, we find that access to the digital service leads to significant improvements in farmers' knowledge and adoption of recommended practices, a modest increase in rice yield and harvest, and a large reduction in the likelihood of rice crop loss on average. Further analyses suggest that the treatment impact is concentrated in areas hit by certain types of weather shocks, increasing harvest by up to 9% and reducing severe crop loss by up to 21% in affected areas. We use vegetation indices (VIs) to construct an objective yield measure for all farmers in the study sample and confirm that our key survey results are robust against differential attrition, reporting biases, and survey sample selection. While the VI-predicted yield provides valuable validation of survey results, our analysis highlights the need for methodological improvements in the effective application of remote sensing data to measure program impacts on agricultural outcomes.

Keywords: Agriculture, digital extension service, climate adaptation, remote sensing
JEL Classification: O12, O13, Q16

*Cole (scole@hbs.edu); Goldberg (jagold@umd.edu); Harigaya (tharigaya@precisiondev.org); Zhu (jzhu@precisiondev.org). We are grateful to the Gates Foundation and the Wellspring Philanthropic Fund for generous financial support, to the Government of Odisha for collaboration and implementation support, to the teams at Precision Development (PxD) and Abdul Latif Jameel Poverty Action Lab (J-PAL), and especially Hannah Timmis, for implementing research activities and providing helpful insights, and to Ofir Reich for his contributions to the analysis of satellite imagery. The study was registered in the AEA RCT Registry: AEARCTR-0008560. International IRB approval was obtained from Harvard Business School. For local ethics reviews, the study was approved by the Institute for Financial Management and Research. Disclosure: Cole is one of co-founders of PxD, but does not receive any financial compensation from the organization. Harigaya and Zhu are employed by PxD. All errors are our own.

1 Introduction

Smallholder farmers represent a significant portion of the global population living in poverty, with many subsisting on less than US$2.15 per person per day (Fu and van Nieuwkoop 2023). These farmers face persistently low agricultural productivity and increasing production risks from climate variability, which adversely affect their livelihoods and contribute to food insecurity. For instance, estimates indicate that maize yields of smallholders in sub-Saharan Africa (Koo 2014) and rice yields in some regions of India range between 25% and 50% of their potential (Das 2012).

While farmers in low- and middle-income countries (L&MICs) have a depth of knowledge passed down by generations, they often do not have access to modern science-based agricultural information (Fabregas, Kremer, and Schilbach 2019) or localized weather information (Fabregas et al. 2023; Rosenzweig and Udry 2019) to optimize their farming decisions. Their adoption of profitable and risk-reducing technologies has remained low, especially in Africa and South Asia (Ashraf, Giné, and Karlan 2009; Duflo, Kremer, and Robinson 2011; Evenson and Gollin 2003; Mobarak and Rosenzweig 2013). For instance, hybrid maize seed varieties can increase average yields and improve resilience to weather shocks but have been adopted by fewer than half of the sub-Saharan African farmers (Bird et al. 2022). A flood-tolerant rice variety has been shown to dramatically reduce the risk of crop loss under submergence and lead to large yield gains in flood-prone areas of Eastern India. Yet the adoption rate remained at merely 10% five years after the introduction of the variety (Janvry, Rao, and Sadoulet 2022).

Agriculture extension services have long been supported by policymakers and practitioners as a strategy to encourage the adoption of modern agricultural inputs and practices. However, it is well-documented that traditional public in-person extension services often lack resources and accountability, thereby limiting their reach and impact (Anderson and Feder 2004; Cole and Fernando 2021; Fabregas et al. 2023). Studies estimate that the farmer-to-extension worker ratio exceeds 1,000:1 in many L&MICs (Fabregas, Kremer, and Schilbach 2019), including in India (Nandi and Nedumaran 2022). With advances in information and communication technologies (ICTs) and the rapid growth of mobile phone ownership, many governments are now integrating ICTs into their agricultural extension programs (Aker 2011; Aker, Ghosh, and Burrell 2016; Fabregas, Kremer, and Schilbach 2019). A growing body of empirical evidence suggests that agricultural extension via mobile phones can increase the adoption of new or unfamiliar technologies, but that their effectiveness varies across products, initiatives, and contexts as one would expect (Abate et al. 2023; Spielman et al. 2021). Importantly, few studies have evaluated the impact of a scaled service reaching diverse farmer populations.

This study evaluates the impact of an agricultural advisory service that provides timely and localized agricultural information via mobile phones to millions of rice farmers in Odisha, India.[1] The service, called Ama Krushi, is a free, two-way voice-based platform that delivers regular audio messages with agricultural advice, and offers an automated hotline that farmers can call to record agricultural questions and receive recorded answers within a few days. Advice is tailored to the data collected during service-registration phone calls, which include farmers' primary crops, geography, and land type, as well as the data collected from hotline reports of other local conditions, such as weather events and field problems.

One persistent challenge in disseminating targeted information and advice to smallholder farmers at scale is the need for local and real-time information to make the content sufficiently relevant for diverse farmer populations. Remote sensing (satellite imagery) would be a potential source of real-time information, but few organizations have the capacity to analyze and monitor such data. Similarly, converting meteorological weather forecasts to agriculture-relevant weather forecasts, and offering specific agricultural practice recommendations, may be too complex a task for many governments or farmer-facing organizations. Because information goods for low-income households are difficult to monetize, existing services tend to be highly customized for a small user base (often offered as fee-for-service by a private service provider or set up as a small-scale program with government or philanthropic funding) or operated at scale with limited scope for customization. The service evaluated in this study was designed to strike the balance where the information content was customized based on a small number of farm characteristics that are easy to collect, and regularly reviewed and dynamically updated by government agronomists based on available information on weather and field conditions.

Our study took place in five districts where the service had not yet been made widely available in 2021. We identified study participants with a random-walk sampling approach, recruiting 13,675 rice farmers across 18 blocks in two phases. This comprised 5,204 farmers before the main agricultural season in 2021 (Cohort 1) and 8,471 farmers before the main agricultural season in 2022 (Cohort 2). We recruited farmers who intended to grow rice and were interested in a phone-based advisory service. In each cohort, 50% of the farmers were

1. The service was developed by a non-profit organization, Precision Development (PxD), in collaboration with the Odisha state government, and and was transferred to the government in 2023 as part of the design.

randomly assigned to the treatment group and invited to register with the digital extension service, while the remaining 50% served as a control group. Primary plot boundaries were collected from all farmers in both cohorts to assess the program impact using satellite-based yields; two rounds of follow-up survey data on agricultural practices and outcomes for the main agricultural season (Kharif) were collected from Cohort 1 farmers.

Three sets of findings emerge from our analysis of survey data from two main agricultural seasons in 2021 and 2022. First, the majority of treated farmers successfully engage with the service: while the proportion of engaged farmers slightly declines from 94% to 84% in the two years, the intensity of service usage among engaged farmers is sustained over time. In the first Kharif season, an average treated farmer picked up 55% of the advisory calls, listened to 52% of content conditional on picking up the call, and spent approximately 28 minutes on the platform. In the second Kharif season, among the 84% of treated farmers who continued to use the service, the average listening rate remained similar at 56%. These patterns suggest a sustained demand for local agricultural advisory in this population.

Second, access to digital advisory improves agricultural knowledge, practices, and production outcomes. We observe improvements in knowledge and adoption behavior across 8-14 practices, with approximately 0.05-0.1 standard deviation increases in the summary indices in both years. This leads to average per-season increases in rice yield (harvest per unit of land) by 1.74% over the control mean and harvest by 4.12% over the control mean. The estimated effect on harvest is robust to multiple hypothesis corrections in the second year and when the data are pooled between the two years. These point estimates are in line with the existing evidence of the impact of digital agricultural information services. For instance, a meta-analysis across seven digital agriculture programs in Asia and Africa shows a 4% increase in yield, on average (Fabregas, Kremer, and Schilbach 2019).

Third, the service helps farmers cope with some (but not all) weather-related shocks. In the second year, in which we collected data on crop losses, treated farmers reported lower likelihoods of experiencing any rice crop loss and severe rice crop loss.[2] These effects are driven primarily by a 10% reduction in the losses that are due to weather-related events, including pests and diseases. We further investigate the impact on resilience to weather shocks by estimating the heterogeneous treatment impact by the presence of major weather events during the evaluation period, which we identified using daily rainfall data and government records on rice field damage due to weather events. While the analysis of heterogeneous treatment effects largely generates imprecise estimates after multiple hypothesis adjustments, three sets

2. We define severe rice crop loss as losing more than half of the rice crops.

of results are worth highlighting. First, we observe consistent, positive effects on agricultural outcomes in areas hit by concentrated, excess rainfall that caused submergence in 40% of study blocks in the first year. Point estimates indicate a 6.8% higher yield and a 9% higher harvest by treated farmers than control farmers in these shock-affected areas. While the increase in yield does not survive multiple hypothesis adjustments, the effect on harvest is statistically significant at 10% after adjustments, and is robust to varying thresholds we use to identify areas hit by excess rainfall. Second, the reduction in severe crop loss in the second year appears to be concentrated in areas that received inadequate rainfall during the growing season, showing a 21% reduction in the likelihood of severe crop loss among treated farmers compared to control farmers in those areas (significant at 1% after multiple hypothesis adjustments). Third, the service does not help farmers cope with every shock. The treatment impact in areas hit by sudden river flooding in the early growing season of the second year is close to zero across all agricultural outcomes.

To shed light on the impact of digital service on farm incomes, we collected data on rice production costs and sales of harvested rice from a subset of farmers in the first year. Agricultural profits are well-known to have a large variance and a distribution with very long tails (Baul et al. 2024; Okorie, Afuecheta, and Nadarajah 2023). This study was not designed to measure profit impact with precision[3], but the analysis could still offer useful insights on plausible impact. Overall, the pattern of treatment impact estimates on profits corroborate the observed effects on total rice harvest: a small, positive point estimate for the full sample and a large point estimate for areas hit by excess rainfall. We take these results as suggestive of increased net profits for treated farmers in areas affected by excess rainfall. Our back-of-the-envelope calculations on cost-effectiveness using these imprecise estimates on profit impact corroborate other studies that show that the marginal benefits of digital advisory services can be an order of magnitude larger than the marginal cost of service delivery (Baul et al. 2024; Fabregas et al. 2023).

A second objective of this study is to explore the potential of using remote-sensing data to measure yields and examine treatment impacts. Key constraints to measuring the impact of interventions on agricultural outcomes include, again, the large variances of outcomes across individuals and seasons, and the reporting biases and attrition issues when using self-reported measurements. A potential solution to these challenges is the use of satellite data, which faces limited to zero attrition, provides more objective measures that are free from reporting

3. In our setting, ex-ante power calculations suggested that we would have needed a survey sample that is an order of magnitude larger to precisely capture the impact on profits.

bias, and offers outcome measurements over a long time horizon (Campolo et al. 2022; Jain et al. 2016; Jain et al. 2019; Lobell et al. 2019). We collected primary rice plot boundaries from all 13,675 farmers in the study, using a GPS device at baseline, to extract vegetation indices; we also collected 1,247 crop-cut (CCE) yields over two agricultural seasons (Kharif 2021 and 2022) as the "ground truth" data. We model the relationship between vegetation indices and CCE yields using machine learning methods, and then use this prediction model to estimate rice yield for the full study sample for the two post-intervention Kharif seasons.

We use the predicted satellite-based yields for three purposes. First, we show that survey results are robust against differential attrition by testing whether the post-intervention yield predicts follow-up survey attrition and whether this correlation differs by treatment status. Second, we show that survey results are not influenced by selection of unobserved characteristics. Estimated treatment effects on VI-predicted yield are comparable across different analysis samples and the full study sample, which addresses the concern that survey results with a 20% attrition may be influenced by selection into completing surveys. Third, we examine whether the VI-predicted yield can measure the program's impact on a farmer's primary rice plot. Using the full study sample, we observe a positive treatment effect on VI-predicted yield in the first year and a null effect in the second year. Further analysis shows that the VI-predicted yield and the self-reported yield generate consistent impact estimates when using the same plot measurement (i.e., GPS-based plot size measured at baseline), but not when self-reported yield is measured using the area of cultivation reported in the follow-up survey. These results identify a new challenge in using remote sensing data to measure program impact on agricultural outcomes for annual crops: one would need updated measurements of cultivation areas across seasons to accurately predict agricultural outcomes over time.

This paper makes contributions to four strands of literature. First, this study expands the evidence base on the impact of digital agricultural extension services in three ways. Many studies in this literature investigate the use of digital tools in increasing the adoption of one or a few agricultural practices. For instance, Dzanku et al. (2020) examine the impact of video documentaries and radio listening clubs on the adoption of an inoculant in Ghana; Arouna et al. (2020) study the impact of a mobile app on the usage of inorganic fertilizer in Nigeria; Fabregas et al. (2025) conduct a meta-analysis of six studies evaluating the impact of SMS messages on the application of lime and fertilizer in East Africa; and Abate et al. (2023) study the impact of a video-mediated extension on three practices for cereal crop cultivation. The service evaluated in this study provides advice on a wide range of agricultural topics and relevant information throughout the course of the season, making the service relevant for potentially more diverse farmer groups. The closest to our study is Cole and Fernando (2021), who examine the impact of a voice-based advisory service — the original service upon which the design of Ama Krushi was based and improved — among cotton farmers in Gujarat, India.

Second, our findings add to a small but growing evidence base on the impact of information services on climate adaptation in L&MICs. An earlier study by Fafchamps and Minten (2012) finds that output market price information and weather forecasts provided via mobile phones to Indian farmers have no impact on their cultivation practices, crop losses, or crop sales. More recently, experimental evaluations on weather forecasts show that precipitation forecasts help farmers optimize labor allocation (Yegbemey, Bensch, and Vance 2023); and improved monsoon onset forecasts influence farmers' crop choice and agricultural investment (Burlig et al. 2024). Our results suggest that a customized advisory service that responds to farmers' diverse information needs could help them cope with adverse weather events, thereby reducing the likelihood of severe crop losses.

Third, only a handful of studies have evaluated a digital agriculture program that has reached scale. Studies included in the meta-analysis conducted by Fabregas et al. (2025) reach several hundred thousand farmers; Baul et al. (2024) examine the impact of a video-based extension program that reaches two million farmers in India. Our study evaluates the impact of a public digital advisory service that was reaching nearly 3 million farmers by the end of the evaluation and is now reaching over 6 million farmers in an Indian state. The key value proposition of digital information services is their ability to reach scale at a low marginal cost. Understanding the impact of these services at scale beyond well-managed research and pilot program settings offers important insights for policy.

Lastly, this paper advances the literature on the use of remote sensing data in impact evaluations (Cole et al. 2025; Jain et al. 2019; Jain 2020; Kubitza et al. 2020). The use of satellites in estimating agricultural yield has been well established in the Global North, but its application in the smallholder context is relatively new (Burke and Lobell 2017; Guo, Chamberlin, and You 2023; Jain et al. 2016; Lobell et al. 2019). We predict rice yield for smallholder farmers in India, whose plots are approximately 0.1 hectare on average, and obtain performance levels that are at par with previous studies with larger plots. While our findings point to the effective use of VI-predicted yield to assess the validity of survey-based impact estimates in an RCT, they also highlight key challenges in using the predicted yield to capture the program impact.

The remainder of this paper is organized as follows. Section 2 provides a general context of the digital agricultural extension service and describes the specific service evaluated in this study. Section 3 describes the setup of the experiment and data used for analysis. Section 4 presents the empirical strategy adopted. Section 5 illustrates the findings from the service data and survey data and Section 6 shows the findings from the remote sensing data. Section 7 discusses the benefit-cost ratio of the digital service and its policy relevance. We conclude in Section 8.

2 Context and Intervention

2.1 Study context

India has been at the forefront of the digital revolution, with its significant investment in developing digital public infrastructure to integrate digital technologies in public-sector services. The most prominent example is Aadhar, the biometric identification system, which streamlined access to public services and welfare programs. In the agricultural sector, the digitization effort was largely decentralized until recently. While a nationwide toll-free call center (called "Kisan Call Center") has been available for real-time advice to farmers who are aware of the service, digital agricultural extension initiatives have been primarily developed and executed by state-level extension offices.

In Odisha, where this study took place, the state government largely relied on the network of public extension agents to disseminate information to farmers, with limited use of digital technologies, before Ama Krushi was launched in 2018. At the time of the study, roughly 4,900 village agricultural workers, agricultural supervisors, and Assistant Agricultural Officers (AAO) served over eight million farmers in the state. This means that the majority of farmers had no interaction with extension agents. In fact, only 25% of farmers in the study sample reported seeking agricultural information from government extension agents, and 5% reported using mobile phones or internet to seek agricultural information at baseline whereas the wireless teledensity of the state was 62.25% in 2020 (Planning and Convergence Department, Directorate of Economics and Statistics 2021). The majority of the farmers identified other farmers and agrodealers as their main information source.

Adoption of modern inputs, and hence yield levels, among rice farmers in the state has remained low. The state average yield of 2,068 kg/ha in 2019-2020 is far below the national average of 2,705 kg/ha (Planning and Convergence Department, Directorate of Economics and Statistics 2021). Additionally, being a coastal state, Odisha is highly prone to natural calamities. Since 2010, the state has been hit by at least one extreme weather event every year, including floods, cyclones, and droughts (Planning and Convergence Department, Directorate of Economics and Statistics 2023). With the high incidences of these weather events, adapting agricultural practices to improve plant resilience to shocks becomes increasingly important.

2.2 Intervention

Ama Krushi is a two-way, voice-based platform that delivers weekly audio messages to farmers and allows them to call in for free, record questions, and receive recorded responses from local agronomists. The service covers a wide range of agricultural topics, from land preparation and seed varietal selection to appropriate application of fertilizer and other inputs, and post-harvest management. Farmers also receive relevant, non-agronomic information, such as government schemes targeting agricultural households and minimum support prices for main agricultural crops.[4]

In addition to standard agronomic advice, three key features of Ama Krushi may help farmers navigate dynamic weather and market conditions. First, the service utilizes publicly available weather information to optimize the timing of advisory messages and provide real-time advice on both preventive and mitigation actions. For instance, advice on sowing is timed to arrive after sufficient rainfall is observed: in cases of delayed monsoon, farmers receive advice to wait until sufficient rainfall is observed. When cyclone alerts are issued by the Indian Meteorology Department (IMD), farmers in at-risk areas receive an alert with advice on how to protect their crops. While farmers can, in theory, access weather information used in this intervention (via the IMD or government website on realized rainfall at the sub-district level), internet access in rural areas is still limited even among smartphone users. More importantly, even if farmers had access to the weather information, it is rarely communicated in a way that is easily usable for agricultural decision-making.

Second, advice on inputs and management practices is customized by land type and sowing method, in addition to geography. For instance, advice on stress-tolerant seeds depends on both farmer location and land type: this is important in areas that face both flood and drought risks.[5] The majority of farmers in Odisha grow nursery plants and transplant, instead of broadcasting rice seeds. Because of the more complicated technology, farmers who

4. The national government sets the minimum support prices for key staple crops every year, and the state government decides how much to procure at these prices.

5. Flood-tolerant seeds are not recommended for farmers cultivating in highland areas; whereas drought-tolerant seeds are not recommended to lowland farmers.

transplant seedlings receive more messages, including advice on the timing of transplanting.

Third, weekly advice is dynamically customized based on the common problems reported via the hotline. If many farmers asked questions about the same problem in a short window of time (e.g., specific pests, excess or scarce rainfall, or market prices), Ama Krushi would push out advisory messages to all farmers in relevant areas about how to tackle that specific problem. This addresses one of the biggest challenges in providing customized agricultural advice at scale: lack of sufficiently local and timely data on agricultural problems faced by farming communities. While the hotline service by itself can partially address this challenge, farmers with low digital literacy may not use the interactive voice response (IVR) system to record questions. In fact, in the first intervention season of this study, only 4% of treated farmers submitted a question to the platform.[6] The ability to broadcast advice based on commonly asked questions (albeit submitted by a small set of farmers who have higher digital literacy) allows temporal and geographic customization of advice for a large number of farmers.

Importantly, Ama Krushi was designed to reach a very large scale at low cost. The advisory was set up as a state government's service in 2017 but initially operated by Precision Development (PxD), a non-profit organization specializing in delivering information services to smallholder farmers at scale. It was fully transitioned to the state government in 2022. At the time of the handover, Ama Krushi was reaching 2.7 million users in the state at the cost of US$0.37 per farmer per year. By the end of 2023, the user base has grown to 6.9 million farmers.

3 Experimental Design and Data

3.1 Study sample and recruitment

To measure the impact of Ama Krushi, five out of 30 districts in the state were set aside from the state-wide rollout plan in the early phase of the program. We listed 21 eligible blocks out of 72 across these five districts, which had a low proportion of rural households already registered for Ama Krushi service and showed low cloud coverage in satellite imagery in the 2019-2020 Kharif season. Within those blocks, we then identified all rural villages that both had more than 50 households according to the 2011 Census and were located in a panchayat

6. The limited usage of the Q&A feature is a common pattern. Based on the platform monitoring data of this digital service, in 2022, only 2.17% farmers who called into the service platform asked a valid agricultural question.

(an administrative unit above the village level) with service penetration below 10%. A total of 1,313 or 39% of the villages met these criteria.

To construct the sample of rice farming households for the study, we used a random walk approach and identified agricultural households that engaged in rice farming, owned a mobile phone, and were interested in receiving a free digital agricultural advisory service. In addition, recruited farmers were asked to provide consent for the enumerator to collect the GPS data on the boundary of their primary rice plot, defined as the largest rice plot within 30 minutes by foot from their house.

Due to the global pandemic in 2020-2021, baseline data collection was interrupted several times, resulting in recruitment of the final sample of farmers in two waves over one year. We recruited 5,204 farmers across 15 blocks into the study before the Kharif season in 2021 (Cohort 1 farmers) and additional 8,471 farmers across 16 blocks later in the same year (Cohort 2 farmers). This resulted in the final study sample of 13,675 farmers across 18 blocks. The timeline of farmer recruitment and other evaluation activities is presented in Figure 1.

3.2 Randomization

Recruited households were randomly assigned to the treatment or control group with equal probability. Randomization was stratified by panchayat, baseline survey version[7], self-reported yield at baseline (above the panchayat median or not), and (for Cohort 2 only) GPS-measured plot size at baseline (above the panchayat median or not). Overall, baseline characteristics between the treatment and control groups are well balanced. Out of 36 tests comparing sample means for the treatment and control groups as shown in Table 1, only two variables for Cohort 1 farmers — primary rice plot yield and having received baseline survey compensation before the intervention — and no variables for Cohort 2 farmers, are significantly different at the 90% confidence level.

The criteria to survey the primary or secondary agricultural decision maker in the household selected through a random-walk sampling approach resulted in a sample predominantly consisting of male farmers who are household heads. The majority are literate and have some formal education. Smartphone ownership was still limited at the time of the baseline survey. Between the two cohorts, 59-67% had a feature phone as their primary phone, and only 40% were the sole owner and user of their primary phone, suggesting that most phones

7. In the baseline survey, half of the sample were randomly selected to be asked about their agriculture knowledge, while the other half were asked about their agricultural practices.

were shared within the household.

Rice is the main crop grown in the Kharif season for nearly all farmers (94-96%), with an average cultivation area of 0.9-1 hectare. Agricultural plots in this region are highly fragmented: farmers cultivate rice on multiple small plots (Das 2012; Rice-based Cropping Systems, n.d.). The primary rice plot is on average 0.14 hectares and accounts for less than 15% of the total cultivation area at baseline. This means that an average farmer cultivates more than five parcels of land. The average of farmers' self-reported total rice yields in the Kharif 2020-2021 season are 2,952 kg/ha and 3,246 kg/ha among Cohort 1 and 2 farmers, respectively, somewhat higher than the state average of 2,730 kg/ha (Planning and Convergence Department, Directorate of Economics and Statistics 2021). The average profits are Rs 10,165 (US$137) for Cohort 1 farmers and Rs 8,366 (US$113) for Cohort 2 farmers — obtained by multiplying the reported harvest with the panchayat-level median sales price minus the reported production cost.

3.3 Data

We use three sets of data to examine the impact of the digital advisory service.

First, we collected three rounds of survey data. The baseline survey was conducted in person with all farmers in both cohorts in 2021; and two rounds of follow-up data were collected from Cohort 1 farmers only. In both follow-up surveys, we randomly assigned whether a farmer was contacted by phone or in-person, and switched survey modality after a specified number of unsuccessful contact attempts in order to maximize response rates.[8] The first (midline) follow-up survey, conducted in early 2022, collected data about the first intervention season (Kharif 2021). We gathered data on agricultural practices, rice cultivation, and harvest from the full sample, and data on agricultural knowledge, rice sales, revenues, and production costs from a subset of farmers that received an in-person survey. The second (endline) follow-up survey, conducted in early 2023, collected data about the second intervention season (Kharif 2022). In addition to the data on rice cultivation and harvest, we asked about the incidents of rice crop loss and reasons for losses. Data on agricultural practices were collected from a subset of farmers that received an in-person survey.

Second, we use the administrative data on service usage to assess how farmers engaged with the service and what types of agricultural information they accessed. The service platform data contains records of all communications with users, including registration survey responses, when farmers picked up advisory calls, for how long they stayed on the line, and

8. Details of the survey modality assignment are described in Appendix B.

whether they called into the hotline and successfully asked questions.

Third, we extract Sentinel-2 satellite imagery between 2020-2022 and obtain vegetation indices for the primary rice plots of all farmers in the study, based on the plot boundaries collected at baseline. These vegetation indices enable us to estimate rice yield for the entire study sample, without attrition, using a yield prediction model which we developed from a separate crop cut dataset collected in 2021 and 2022 from 1,247 farmers in the study blocks.

3.4 Key outcomes

The main outcomes of interest include farmers' agricultural knowledge, adoption of recommended practices, and rice crop production. This study was designed to detect 2-3 percentage point improvements in agricultural knowledge and adoption, using the survey data from 5,204 Cohort 1 farmers, and a 3 percent increase in yield, using the satellite-based yield of the full sample of 13,675 farmers.

To measure agricultural knowledge and adoption of recommended practices, we asked a local agronomist, in consultation with experts from agricultural research institutions, to identify key practices before the intervention, and we collected data on those individual practices at each survey round.[9] We then create inverse covariance weighted summary indices, following Anderson (2008), by aggregating knowledge responses for the knowledge index and reported behaviors for the adoption index. To measure yield and harvest, we use self-reported production and area of cultivation in each survey round. Agricultural outcomes are known to be noisy and have a distribution with very long tails (Baul et al. 2024; Okorie, Afuecheta, and Nadarajah 2023).

Additionally, we construct measures of profits from data on production costs and rice sales, that we gathered from a subset of farmers who received an in-person survey at midline. The average farmer in our sample sells 35% of harvested rice; the rest is probably stored and consumed by the household over time. Our primary profit measure is imputed, to include revenue from sales and the value of the unsold harvest at the panchayat-median sales price, from which we subtract the total spending on variable production cost items (as outlined in the preanalysis plan, PAP); we also report results using two alternative measures in the appendix. The first of these alternatives improves the precision of the profit measure by taking winsorized cost items (instead of raw values) to create a measure of harvest cost. As discussed later, this improves the precision of impact estimates without substantially influencing point estimates. The second alternative values the unsold harvest at the retail

9. The detailed description of each practice is listed in the Appendix Table A3.

price of rice in Odisha in 2021, or Rs 29 per kg, based on publicly available price data (Consumer Affairs, n.d.). The median sales price in our survey data is Rs 14 per kg, or 50% lower than the retail price, thus leading to meaningful differences in impact estimates.

We winsorize non-negative continuous variables, including the cultivation area, yield, harvest, value of harvest, harvest cost and investment, at the 95th percentile, and the profit measure at the 2.5th and 97.5th percentiles, to mitigate the influence of extreme outliers.

3.5 Survey attrition

A key concern in the analysis of survey data is the influence of survey attrition on estimates of the treatment impact. There are two relevant issues. First, different types of farmers may complete follow-up surveys between the experimental groups, resulting in biased treatment impact estimates. Second, even in the absence of differential attrition by the experimental group, farmers who respond to follow-up surveys may be systematically different from those who do not. The comparison between the experimental groups still generates unbiased impact estimates for those who complete the survey, but these estimates may be different from the average impact of the full study sample population.

In Appendix Table A1, we show that 80% of Cohort 1 farmers completed at least one of the follow-up surveys, and 66% completed both. Survey attrition is correlated with farmer characteristics, but there is no differential attrition by treatment status. P-values of the joint F-test across interactions between the treatment status and baseline characteristics are 0.495 and 0.228 for the midline and endline surveys, respectively. We investigate the second issue — whether the selection into the survey sample results in impact estimates that are substantially different from the average impact of the study sample in Section 6.3 — by comparing impacts estimated using VI-predicted yield from the analysis sample and the full study sample.

4 Empirical Strategy

We registered a PAP prior to the collection of the first follow-up survey.[10] The empirical strategies described below largely follow our PAP.[11]

We estimate intent-to-treat (ITT) effects for all key outcomes. When baseline values are available, we use the ANCOVA specification; otherwise, we use the OLS specification. For

10. The study is registered in the AEA RCT Registry: AEARCTR-0008560.

11. We show additional analyses, such as robustness checks, that are described in the PAP but not presented in the main paper in an online appendix.

outcomes collected in both rounds of follow-up surveys, we estimate the effects separately for each year and also the average effect over two years using the pooled data to increase statistical power. Specifically, our main specification is as follows:

$$Y_{it} = \alpha + \beta T_i + \psi Y_{i0} + X'_{i0}\delta + S'_{i0}\eta + \gamma_s + \epsilon_{it}, \tag{1}$$

where Y_{it} stands for outcomes of interest of farmer i in year $t \in (1, 2)$, with $t = 1$ and 2 indicating the Year 1 (Kharif 2021) season and the Year 2 (Kharif 2022) season, respectively; T_i is the treatment indicator; Y_{i0} is the pre-intervention value of the outcome measured at baseline; X_{i0} is a vector of exogenous control variables listed in Table 1; S_{i0} is a vector of survey design variations including indicators for the survey modality, whether the final survey modality is switched from the initial assignment, and the timing of follow-up data collection; and γ_s represents the stratification block fixed effects. We use Huber-White robust standard errors in analysis using a single round of follow-up data, and cluster standard errors at the farmer level in analysis with the pooled data. To prevent loss of power from missing values in the baseline outcomes and covariates, we substitute those missing values with the median value within each block and cohort, and include an indicator for imputed values, as pre-specified in the PAP.

We are also interested in exploring heterogeneous treatment effects. We estimate:

$$Y_{it} = \alpha + \beta_1 T_i + \beta_2 T_i Z_i + \psi Y_{i0} + X'_{i0}\delta + S'_{i0}\eta + \gamma_s + \epsilon_{it}, \tag{2}$$

where $T_i Z_i$ represents the interaction between the treatment indicator and the subgroup variable of interest; other notation is the same as in Equation 1.[12] Specifically, we consider heterogeneous impact by agricultural yield measured at baseline and the weather shocks identified ex-post using public weather data and government reports.

We correct for the multiple hypothesis testing within each category of key outcomes – rice production, rice income, rice crop loss, and heterogeneity by weather shocks. We adjust for the family-wise error rate using the bootstrap resampling method (Romano and Wolf 2016) and report both the "naive" p-values and Romano-Wolf step-down adjusted p-values for specifications reported within the same table. We do not control for multiple hypothesis testing for secondary outcomes, including adoption of individual practices, specific rice production cost and investment items, and primary rice plot cultivation, as these represent

12. Because our subgroup variables of interest are all perfectly collinear with the stratification block fixed effects, we do not control for Z_i in this specification.

exploratory analyses that aim to provide qualitative insights.

5 Empirical Results from Service Platform Data and Survey Data

5.1 Engagement with the advisory service

We start by summarizing the uptake and the usage of the Ama Krushi advisory service. Figure 2 plots kernel densities of the service engagement of Cohort 1 farmers over two Kharif seasons. Consistent with statistics reported in Tables 2 and 3, nearly all farmers who were assigned to the treatment group registered for the advisory service, and the majority of registered farmers engaged with the service and accessed agricultural content. The proportion of farmers that engaged in the service declined between the two years from 94% to 84% of Cohort 1 treated farmers, but the engagement level among those who used the service was sustained. In the first year, as reported in Column (1) of Table 2, an average Cohort 1 farmer spent a total of 27 minutes on the advisory platform, picking up 55% of 52 advisory push calls and listening to the majority of the content in 24% of calls. In the second year, Ama Krushi sent shorter calls in general, resulting in a 40% decline in the average time spent on the service platform among Cohort 1 farmers, from 27 to 16 minutes. Farmers listened to the majority of 21% of these calls, about the same as in the first year.

We note that treated farmers who consented to take follow-up surveys tend to be more engaged with the service than those who did not (Column (5) in Table 2 and Column (7) in Table 3). Since we do not observe engagement for the farmers in the control group, we cannot test whether the treatment and control samples are balanced on engagement with (or propensity to engage with) the service. If the farmers in the control group who participate in the follow-up surveys were to have a higher propensity to engage with the service than the control farmers who attrit from the sample, then our analysis sample provides an unbiased estimate of the effect of the service on a selective sample of farmers with a higher propensity to engage. While this impact is unbiased, it may be larger than the impact on the average farmer population, which we examine further in section 6.2. If, however, the control group farmers are comparable to the treated farmers on observables but different in terms unobserved propensity to engage with the service, then estimated impacts could be biased up.

5.2 Intention to treat effects

We now turn to the average treatment impact on farmer behavior and agricultural outcomes using the ITT analysis. The results reported in this section use the two rounds of follow-up survey data collected from Cohort 1 farmers.[13]

5.2.1 Impact on knowledge and adoption of recommended practices

Table 4 reports the ITT effects on knowledge (measured in Year 1 only) and adoption of the core index — a subset of practices that are relevant for all farmers regardless of their planting method — and the transplanting index — a broader set of practices that are relevant for farmers who practice transplanting. Both indices are standardized with a mean value of zero and a standard deviation of one in the control group. When estimating effects on the core index, we limit the analysis sample to farmers who planted rice in a given season, and when studying the transplanting index, we further restrict to farmers who reported transplanting rice at baseline (roughly 85% of Cohort 1 farmers). Overall, treatment has robust, positive effects on farmer's knowledge and self-reported adoption of recommended practices in the first year. Access to the digital advisory service improves agricultural knowledge by 0.11 SD and agricultural practices by 0.068-0.098 SD. In the second year, the average impact on agricultural practices as measured by the core index becomes somewhat smaller (0.050) and not significantly different from zero at conventional levels, while the point estimate is similarly large and precisely estimated for agricultural practices measured by the transplanting index (0.111) despite the smaller sample size in the second year.

Appendix Tables A4 and A5 report the treatment effects on the adoption of individual agricultural practices. Consistent with the existing evidence of the impact of digital advice on farmer behavior (Fabregas et al. 2025), we observe increases of a few percentage points each across several practices in both years, including seed replacement rate, pesticide adoption, and zinc application in Year 1, and seed replacement rate, and nursery fertilizer application in Year 2.

5.2.2 Impact on agricultural outputs

We next examine treatment impacts on cultivation decisions and agricultural outputs. Table 5 Columns (1) and (2) show the ITT effects on decisions about whether to plant rice and about the land allocated to cultivation of rice; Columns (3) and (4) report the ITT effects on yield (measured in kg/ha) and harvest (measured in kg). We report impact estimates

13. Cohort 1 treated farmers started receiving agricultural advisory service on June 9, 2021 (the Kharif 2021 season).

separately for each year (Panels A and B) and together in a pooled regression (Panel C). Three insights emerge from this analysis.

First, we find no evidence that access to digital advisory affects rice cultivation decisions for an average farmer, at either the extensive or the intensive margin. In both years, more than 95% of farmers in the sample report planting rice with an average cultivation area of a little less than one hectare.

Second, impact estimates on yield and harvest suggest improved agricultural outputs for treated farmers. The point estimates of effects on both yield and harvest are consistently positive, and precision improves in the pooled sample. In Figure 3, we plot kernel densities of yield across all rice plots by experimental arm. The density curves are indistinguishable between the treatment and control groups at baseline (Sub-figure a). In contrast, the density curves for the treatment group shift right in both post-intervention years (Sub-figures b and c), suggesting improved yield over a large portion of the distribution.[14]

Third, the estimated effects are noticeably larger and more precise for harvest than for yield. Pooling data from both seasons, treatment increased yield by 1.7% (Column (3) in Panel C). Despite no significant increase in the area dedicated to rice cultivation, treatment increased the quantity of rice harvested by 4.1% relative to the control mean of 2,468 kg. These results are robust to multiple hypothesis correction.

It is plausible that the average effect masks important heterogeneity in the impacts on cultivation areas and yield. For instance, digital advisory may encourage farmers with resource constraints to intensify input use in a smaller cultivation area, while it may encourage others to expand areas of cultivation. We explore potential mechanisms further in the later discussion of heterogeneous impact by baseline yield in subsection 5.3.3.

5.2.3 Impact on revenues, costs, and profits

In Table 6, we report estimated effects on revenues, value of harvest, costs, and profits from rice production. Recall that we only collected these data from farmers who completed in-person surveys: harvest data, on the other hand, were collected in both in-person and phone surveys. Therefore, we begin by reporting the treatment effect on harvest in the full sample of 3,835 farmers: an increase of 76.5 kg, representing a 3.22% change relative to the control group. We compare this to an estimated effect of 108.0 kg (4.82%) for the subset of 1,929 farmers in the in-person sample. We fail to reject the equality of the treatment effect among

14. Two-sample Kolmogorov-Smirnov tests show that the the difference in yield distributions between the control group and the treatment group are statistically insignificant in Year 1 and significant in Year 2.

these two samples. We also fail to reject that farmers assigned to be surveyed in person have the same yield as those assigned to be surveyed over the phone when using satellite data (Table 12), suggesting no systematic difference between the two samples.

Columns (3)-(6) of Table 6 report the treatment effects on revenue, cost, and profit outcomes, for the subset of farmers who completed in-person surveys. The estimates are positive across all outcomes but imprecisely estimated. The coefficients represent increases of 3.95% in reported revenues, 4.44% in the value of the harvest, 3.20% in variable costs, and 4.24% in an imputed measure of profits.[15] The confidence interval for the estimated effect on profits is particularly large with the possible effects ranging from a decline of Rs 1,493 to an increase of Rs 2,146. The treatment impact estimates for alternative construction of outcome variables using winsorized cost and different revenue components are reported in Appendix Table A8, Columns (5)-(6). The point estimates remain relatively consistent, while the precision of the estimates improves as expected. Still, with the wide confidence intervals, these average impact estimates do not offer clear evidence of the impact on economic outcomes for farmers.

5.2.4 Impact on likelihood of crop loss

In the second year, we collected data about rice crop losses in order to understand whether the treatment improves a farmer's ability to cope with adverse events. Table 7 presents the treatment impact on the likelihood of any rice crop loss and of severe rice crop loss (defined as losing more than half of their crops), broken down by reported reasons for loss. Notice that the incidence of crop loss is high in this population: over 61% of control farmers reported experiencing some rice crop loss (Column (1)), and 21% of control farmers lose more than half of their crops (Column (6)). Access to Ama Krushi reduces the likelihood of any crop loss by 4.9% and of severe crop loss by 9.6% (Columns (1) and (6)). Most of the severe crop loss is caused by weather-related events, including floods, other weather events, and pests and diseases. Treatment reduces the probability of any loss due to other weather by 3 percentage points; effects on other channels of loss are estimated imprecisely. Focusing on the probability of severe loss, treatment sharply reduces the probability of loss due to other weather events, and to pests and disease, by 24% and 26% relative to the control means. These results are robust to multiple hypothesis corrections. Unfortunately, we did not collect details of "other weather events" that farmers reported. Among a small group of farmers who provided additional information (4% of those that reported "other weather events"),

15. We report the treatment effects on individual rice production cost and investment items in Appendix Table A6. The treatment triggers some variable cost (e.g., on fertilizer, hired labor, and tractor) and increases the investment amounts.

excessive water (i.e., heavy or unseasonable rainfall that did not cause floods) and drought or scarcity of water are the two commonly reported reasons.

There are several features of Ama Krushi that may help farmers manage weather risks. First, the advisory service provides prophylactic advice to improve plant resilience. For instance, stress-tolerant seed varieties, including drought- and flood-resistant varieties, are recommended in at-risk areas; zinc fertilizer can improve plant performance under water-stress conditions; and preventive pest management practices reduce the potential damage from pest attacks. Second, reactive advice to control damage from extreme weather events is sent in real time: The timing of advice on sowing is based on realized rainfall; farmers receive advice on the re-application of fertilizer and pest management after heavy rainfall in the early growing season. Finally, farmers receive alerts and relevant advice when cyclones are forecasted, including advice to cover crops during the growing season and harvest early for storage in a dry place during the harvest season.

Notably, reported severe crop losses due to floods in 2022 are concentrated in areas that were affected by a sudden overflow of the Mahanandi river in August, an early growing season. This event resulted in the flooding of a large proportion of rice fields in the villages located along the river bank [16]. In these areas, 45% of farmers in the control group reported losing more than half of their rice crops, 85% of which were reportedly due to floods. In other areas, only 2.3% of control farmers reported losing more than half of their rice crops due to floods. The advisory service did not have the information in advance to alert farmers of the flooding. Even if the information were available a few days in advance, there may be limited real-time advice that could help farmers mitigate the effect of a severe flooding event. This may be an example of a weather shock that digital advisory alone cannot help farmers to protect themselves against.

5.3 Heterogeneous treatment impacts

The analysis of the average treatment impact suggests that Ama Krushi significantly improves agricultural practices, leading to a modest increase in yield and harvest, and a significant reduction in the likelihood of crop loss due to weather-related events. While the positive point estimates suggest a potential increase in the profitability of rice cultivation, the average effect is estimated imprecisely. In this section, we consider heterogeneous treatment impacts by weather shocks during the evaluation period and baseline agricultural productivity.

16. More detailed information about this river flooding is described in this Wikipedia page: "2022 Odisha Floods".

5.3.1 Extreme weather events during the study period

The reduced likelihood of severe crop loss due to weather-related events suggests that Ama Krushi was particularly effective at helping farmers cope with certain types of shocks. To better understand the extent of this impact, we investigate the treatment impact in the presence of unanticipated weather shocks. To do this, we first identify major weather events that affected rice cultivation in the study area, using daily rainfall data at the block level from the state weather monitoring portal and government reports on damages in rice cultivation from extreme weather events.[17] The identified events are summarized in Appendix Table A7. In the first year, study areas experienced two events: excess rainfall (heavy, concentrated rainfall that caused submergence) during the growing stage[18], and a cyclone in early December at the time of harvest[19]. These events affected 30% and 46% of farmers, respectively, in the study areas. In the second year, there were no incidents of excess rainfall or cyclone in the study areas. Instead, two major events were river flooding (discussed above)[20], affecting 17% of farmers in the sample, and scarcity of rainfall that resulted in a consistent lack of water in rice fields during the growing season, affecting over 33% of farmers in the sample[21].

Excess rainfall in 2021 and scarce rainfall in 2022 are apparent in the rainfall patterns over the course of the season, aggregated at the district-level in Figure 4. In the figure on the left, the spike in the volume of rainfall within a 2-week time window in early September reflects the concentrated heavy rainfall received over two days. In contrast, the figure on the right illustrates the consistent lack of rainfall observed in some blocks over the course of the growing season in 2022.

17. The daily rainfall data are extracted from the government's weather monitoring portal and government's reports on damages on rice cultivation from climate events are extracted from Special Relief Commissioner.

18. We define excess rainfall as receiving more than 250 mm rainfall over two days in the period of August 16 to October 15 in 2021. Shifting the threshold for excess rainfall slightly (50 mm) upward or downward does not affect the overall results.

19. We define a cyclone as receiving more than 50 mm rainfall over two days in the period of December 4 to December 10 in 2021, around the peak harvest time. This definition aims to catch the impact of Cyclonic Storm Jawad.

20. We define river flooding, using the government's incidence report which presents the extent of damage in rice fields at the village level. We aggregate this data to the panchayat level to match our randomization stratification clusters.

21. We define scarce rainfall as receiving less than 450 mm rainfall over the period of August 16 to October 15 in 2022. Shifting the threshold for scarce rainfall slightly (50 mm) upward or downward does not affect the overall results.

5.3.2 Heterogeneous treatment impact by weather shocks

Tables 8 and 9 report the heterogeneous treatment impact by each weather shock. In Year 1, treatment has positive effects on yield, harvest, and profits for farmers living in areas affected by excess rainfall, as reported in Table 8, Panel A. Farmers in these rain-affected areas who were offered Ama Krushi had 165.5 kg/ha (6.8%) higher yields, 198.5 kg (9.4%) higher harvest, and Rs 2,240 (US$30) more in profits than control farmers in the same areas. Only the effect on harvest remains statistically significant after multiple hypothesis corrections accounting for the 36 coefficients reported in Table 8 and 9. When constructing the profit measure using winsorized cost components, the point estimate is similar but the Romano-Wolf p-value reduces to 0.202, as reported in Appendix Table A9 Column (1). The effect of the treatment for farmers in cyclone-affected area is similar, though the benefits are smaller and not statistically different from zero.

In the second year, the characteristics of the weather shocks and the corresponding treatment effects appear to differ from those observed in the first year. Table 9 shows that farmers in areas that were *not* affected by river flooding generally saw the positive effects of the Ama Krushi service, with increases in yield and harvest and a decrease in the probability of severe crop loss; these effects lose precision after correcting for multiple hypothesis testing. However, the effect of treatment in the areas that were affected by river flooding were close to zero. This may be because there are few available options to mitigate against the harm of flooding: none of the advice that Ama Krushi offered to farmers could prevent losses from this type of shock. Treatment had about equal (small, positive, and imprecise) effects on yield and harvest for farmers in areas that did and did not suffer from rain scarcity. Treatment did significantly reduce the probability of severe rice crop loss in scarce rainfall areas, by 5.8 percentage points or 21% of the mean of control farmers in those areas.

As we noted, this study was not designed *ex-ante* to measure heterogeneous impact by specific types of weather shocks. The consistent results between the average ITT impact and heterogeneous treatment impacts by weather shocks provide suggestive evidence that the impact of Ama Krushi is largely concentrated in areas that experienced certain weather shocks.

There are four mechanisms through which Ama Krushi could have had stronger effects on the agricultural outcomes of farmers affected by weather shocks. First, precautionary advice could have made treated farmers more resilient, such as by using more stress-tolerant seeds or improved planting and drainage practices. Second, real-time and ex post advice tailored to realized shocks could have helped treated farmers mitigate damage and recover more effectively than those in the control group. Third, there could have been a higher return to (or take-up of) general agricultural advice in places that experienced shocks than in places that did not. We can only provide indirect tests of these mechanisms, because we did not collect data about specific practices associated with loss prevention and recovery (i.e., we did not ask about the specific seed type used by farmers in the sample).

The precautionary channel cannot explain the effects for farmers affected by excess rainfall in the first year, as Ama Krushi did not issue precautionary advice about rainfall or flooding before the heavy rainfall and cyclone occurred. Farmers in shock-affected areas were more likely to pick up calls about storm mitigation than farmers in unaffected areas, which would be consistent with the second channel, but conditional on picking up the call, farmers in shock-affected areas did not listen to the message at higher rates.

In the second year, the service did offer two types of precautionary advice that turned out to be relevant: advice about planting flood tolerant and drought tolerant seeds. While we do not observe specific planting practices that would allow us to confirm that farmers adopted this advice and that the improved seeds were protective, the precautionary channel cannot be ruled out in the second year. Listening patterns after the second year shocks were similar to those in the first year: shock-affected farmers were more likely to pick up calls that included mitigation advice but not more likely to listen to the content than farmers not exposed to the weather shocks, conditional on picking up the call. The return to this advice could have been higher in shock affected areas, contributing to the higher overall effect of the service for farmers exposed to scarce rainfall.

5.3.3 Heterogeneous treatment impact by baseline yield

One important consideration when assessing the impact of a digital intervention targeting low-income households is how the benefits are distributed across the population. In particular, improved agricultural advice may have different impact for farmers operating at different yield levels. Following the PAP, we report in Table 10 the heterogeneous treatment impact by whether the self-reported yield at baseline is above or below the panchayat median.[22] Differences in the control group means (shown at the bottom of the table) illustrate that farmers with below-median yield at baseline utilize fewer recommended inputs and practices, have smaller cultivation areas, produce and earn less, and are more exposed to the risk of severe crop loss. For this group, access to digital advisory significantly increases areas

22. Yield at baseline refers to the self-reported yield of the primary rice plot for the Kharif 2020 season. We stratified the intervention randomization by this same variable.

of cultivation and harvest (Columns (2), (4), (7), and (9)) in both years, with small and imprecisely estimated positive coefficients for adoption of core agricultural practices, yield, and profits as well. The increase in the area cultivated appears to be the most important mechanism for increased harvest for the low-yield baseline group: the magnitude of the effect on the area cultivated could explain 60% of the increase total harvest, even at the low-yield control group's baseline yield.[23]

On the other hand, treatment impact among farmers with baseline yield above median are mixed. The treatment improves the adoption of core agricultural practices by 0.09 SD and results in a positive and insignificant increase in yield in Year 1, but this does not lead to improvements in harvest or profit. In Year 2, we observe positive but insignificant increases in yield and harvest, leading to a lower probability of severe rice crop loss, but without improvements in measured agricultural practices.[24] When pooling the data between the two years, the treatment effect on yield becomes marginally significant, suggesting that the impact pathway for this group is probably through increased yield (results are not shown but available upon request). Overall, these findings suggest that different groups of farmers may respond differently to the expanded availability of localized agricultural advice.

6 Empirical Results from Remote Sensing Data

Measuring agricultural yield remains a key empirical challenge of evaluating the impact of agricultural programs. Self-reported data are well-known for measurement errors (Abay et al. 2019; Abay 2020; Deininger et al. 2012; Gourlay, Kilic, and Lobell 2019). Yields collected through crop cut exercises are considered to be the gold-standard but are costly. Agricultural outcomes exhibit a large, unpredictable variation across seasons (Rosenzweig and Udry 2019). Multiple seasons of data collection are recommended to allow more precise and robust impact analysis.

A promising solution to these challenges is the use of remote sensing data, which provides objective yield measures that are free from reporting bias, faces limited to zero attrition, and potentially offers outcome measurements for a large sample and over a long time horizon.

In this section, we explore the application of remote sensing data in assessing the impact of the digital advisory service, and discuss its advantages and shortcomings. We first describe

23. In the first year, the mean yield in the control group (2,569 kg), multiplied by the treatment impact on cultivation area (0.036 ha) is roughly 59% of the treatment impact on harvest (157 kg).

24. The total effect of treatment on this group is computed as the sum of the coefficients on treated and the interaction term, and is reported in the bottom row of the table.

the approach to yield prediction and then present two applications of the VI-predicted yield measure: investigating survey attrition and selection, and estimating the program impact.

6.1 Predicting rice yield using satellite imagery

We provide a brief description of the data and our approach to predicting rice yield in this section.

Our vegetation indices come from Sentinel-2 satellite imagery collected between June and December over three years (2020-2022) for the primary plot locations identified at baseline. Satellite imageries are recorded every five days in each orbit. We first create a set of vegetation indices (VIs) — NDVI, reNDVI, GCVI, Laigreen, EVI, and MTCI — at the pixel-day level and aggregate the VIs to the plot-day level using an area-weighted average. We then construct a set of VI features to be used in the yield-prediction model, including monthly and seasonal mean, median, minimum, and maximum values of each VI. In addition, we gather other covariates that are probably correlated with rice yields, including geospatial and climate variables. Lastly, we create an additional set of prediction features by interacting VI features with block dummies to allow the slope of VIs to vary by location.

Our ground truth comes from the CCEs we conducted in 2021 and 2022. In the first year, we collected crop-cut data from 733 farmers who lived in the study blocks but did not belong to the study sample; in the second year, we collected crop-cut data from 150 Cohort 1 farmers, 188 Cohort 2 farmers, and 176 farmers of the first-year CCE sample.

We use several prediction models — OLS, Lasso, random forest, gradient boosting, neural network, and stacking — on various iterations of our prediction feature dataset. These iterations include variations in the threshold for cloud coverage, construction of VI features, and pixel selection along plot boundaries. This approach allows the optimal model to be identified empirically: we conduct a large number of prediction models and select the best-performing model using predetermined performance criteria evaluated at the testing sample with cross-validation. We use two performance metrics. First, the root mean square error (RMSE) is a common metric that measures the average difference between predicted and actual values, thus indicating how well the model can predict the target value (Klompenburg, Kassahun, and Catal 2020). Second, in an impact evaluation, we care about how well the change in predicted yield correlates with the observed change in yield. We thus use the R^2 from a regression that models the linear relationship between the *change* in the predicted yield values and the *change* in crop-cut yields between 2021 and 2022, as the second performance metric. We rank the prediction models by the two performance metrics and select the one with the best overall performance as our optimal prediction model.

Our selected, optimal prediction model uses a random forest method, resulting in an RMSE of 1,194 kg/ha and an R^2 of 0.6122 in the testing data. We highlight three caveats in interpreting the impact estimates using the predicted yield measure we constructed. First, the mean yield in the training sample is 3,227 kg/ha. The observed RMSE in our model is not trivial, posing some concerns about the reliability of the predicted yield measure. Second, the distribution of predicted yield values has a compressed distribution compared to that of CCE yields, as shown in the Appendix Figure A2. This problem has been widely documented and discussed in the literature (Jain et al. 2016; Lobell et al. 2019). In the context of an impact evaluation, the compressed yield prediction probably leads to significantly smaller impact estimates than the true impacts of the intervention in absolute terms. Third, our training data contained a limited number of observations from the plots that reported crop loss, potentially limiting our ability to accurately estimate yield for low-performing plots. In the crop cut sample, 69% of the farmers in the first year and 24% in the second year reported experiencing field issues whereas 61% of Cohort 1 farmers reported experiencing rice crop loss in the second year. While these challenges hinder our ability to rely on VI-predicted yield to accurately estimate program impact, our analysis that focuses on the consistency of patterns in impact estimates, with the above caveats in mind, still offers valuable insights.

6.2 Investigating survey attrition and selection using predicted yield

One major advantage of using satellites to measure yield is that it eliminates attrition in post-intervention measurements. Using predicted yield measures available for all farmers in the evaluation sample, we first revisit issues around survey attrition. In Table A2, we test whether the post-intervention, VI-predicted yield predicts survey attrition and whether this correlation differs by the treatment status within the sample of farmers assigned to receive an in-person survey (Column (1)) and for the full sample (Columns (2)-(4)).[25] The outcome in Columns (1)-(2) is an indicator for responding to questions on agricultural profits in the midline survey (Panel A) and responding to questions on agricultural practices in the endline survey (Panel B). Across the board, we see neither correlations between VI-predicted yield and the likelihood of responding to key survey questions, nor differential correlations by the treatment status. The survey response rates estimated at the mean value of post-intervention, VI-predicted yields are statistically indistinguishable between the control and treatment groups across all comparisons except one. The mean predicted yield among those

25. We calculated bootstrap standard errors for all analyses using VI-predicted yield.

who responded to questions on agricultural practices is 3% higher in the treatment group than in the control group. These results, along with the absence of differential attrition by observable characteristics between the treatment and control groups reported in Appendix Table A1, offer robust evidence that treatment effects estimated using survey data are not influenced by differential attrition.

We next use VI-predicted yields to examine whether the (non-differential) selection into the survey sample influences the treatment impacts estimated with the survey data. We estimate the average treatment impact on VI-predicted yield across different analysis samples and full samples in Table 12. The average effects are generally comparable across different sample frames in both years. Point estimates appear to be slightly larger for some of the analysis samples in the first year (Columns (2)-(3) vs Column (4) in Panel A), but the confidence intervals largely overlap. The small difference may be partially driven by likely survey attrition among farmers who did not cultivate rice. These results suggest that the treatment impacts analyzed using the survey sample are not systematically different from treatment impacts estimated from the full study sample, in one of the key outcomes. It reduces the concerns on the generalizability of treatment impacts estimated from the survey data, given we only follow up with Cohort 1 farmers and experience about 20% attrition in each survey round.

6.3 Estimating treatment impacts using predicted yield

We examine the treatment impact on farmers' cultivation of the primary rice plot, for a subset of farmers who respond to primary rice plot questions, shown in Table 11. The results of interest are the treatment impacts on self-reported yield using cultivation area reported in follow-up surveys (Column (4)), self-reported yield using GPS-measured plot size at baseline (Column (5)), and the VI-predicted yield for the GPS-measured plot (Column (6)). In the first year, the direction of treatment effects on the self-reported and VI-predicted yield are all consistent and positive. The point estimate from the VI-predicted yield is smaller but more precise, potentially because of the issue of compressed distribution of predicted yield discussed above. In the second year, despite there being no treatment impacts on the cultivation areas, the treatment impact estimated from survey data are inconsistent between two yield outcomes that use different measurements of plot area. The treatment impact on the VI-predicted yield is consistent with the effects obtained for the reported yield with the GPS-measured area. One plausible explanation for these patterns is that farmers' rice cultivation areas change across seasons. The GPS-measured plot sizes at baseline were collected from March to May 2021 for Cohort 1 farmers, immediately before the start of the agricultural season, likely representing a more accurate measurement of the cultivation area for the agricultural season in the first year than for the season in the second year. This offers an important insight on the study design when using satellites for yield measurement, in that repeated measurements of cultivation areas may be necessary to use VI-predicted yield to measure program impact over multiple seasons in a setting where farmers cultivate different portions of the plot across seasons. The consistency between effects estimated from the VI-predicted yield and from the reported yield using the GPS-measured area suggest that using remote sensing data to measure treatment impacts has potential.

Taking the advantage of VI-predicted yield is available for the full study sample, we examine the treatment impacts in these large samples. As shown in Table 12, the treatment impact on the VI-predicted yield of the full samples is similar to the impact of the smaller survey sample. On average, the treatment has modest but positive effects on agricultural yield in the first year, and no meaningful effects in the second year.

7 Benefit-Cost Ratios

Digital information services can be delivered at scale at a very low marginal cost per user served. These services can, therefore, generate large benefit-cost ratios (BCRs) at scale even when they have relatively modest average effects on agricultural profits among all farmers reached by the service. To shed light on how marginal BCRs change over time, we calculate BCRs under different time frames at varying levels of scale and prevalence of weather shocks. We use the estimated farmer-level impacts in 2021 from this study, the observed scale and cost of the Ama Krushi service between 2021 and 2024, and weather data over the last 10 years to conduct this exercise.

Recall that we estimated the treatment impact on profits in 2021 using two specifications: the average impact for the full sample and the heterogeneous impact by the presence of excess rainfall. Our earlier analysis suggests that the impact of Ama Krushi is concentrated in areas hit by excess rainfall. In this exercise, we start by presenting the BCR calculations for 2021 using average impacts, and again, focusing on the larger impacts concentrated among the smaller number of farmers affected by excess rainfall, and then we expand the calculations by varying the parameters on scale (farmer reach), cost, and the prevalence of weather shocks using the impact on profits for areas hit by excess rainfall, which are more precisely estimated than the average impact for the full sample.

Panels A and B of Table 13 present the BCRs in 2021 using the average impact on profits and the impact for farmers in areas hit by excess rainfall, using treatment effects as estimated from the experiment and the actual coverage and cost data that correspond to the experimental period. Farmer reach (Column (1)) captures the number of farmers served by the service in a given time period by taking the weighted average of the total number of Ama Krushi users recorded on a quarterly basis. For instance, the number of farmers served in 2021 is calculated by taking the weighted average of farmer reach figures reported in March, June, September, and December in 2021. Ama Krushi served 1.37 million farmers in 2021, 21% of whom resided in areas affected by excess rainfall based on the block-level daily rainfall data for the entire state. The total cost of the service delivery reflects the total budget of the program. In 2021, Ama Krushi was financed by a combination of philanthropic funding and the state government budget at nearly US$1 million or US$0.76 per farmer per year. The resulting BCRs are reported in Column (5). The central estimates across different specifications and profit measurements are all well above 1:1, but the 95% confidence intervals (CIs) vary widely, from large negative to large positive figures. As expected, CIs for the BCRs calculated using more precise profit impact estimates for areas hit by excess rainfall are substantially tighter, with the most precise estimate suggesting a BCR of 10:1 with a 95% CI of 3:1 to 17:1. Notably, the central BCR estimates using the least noisy profit measure, which is based on winsorized cost components, are similar in magnitude between Panels A and B.

In the next three panels, we show how the BCR estimates change from year to year and over time, depending on the scale and the prevalence of weather shocks. Excess rainfall shocks were much more prevalent in 2021 than in the ten-year period 2014-2023. Therefore, we consider whether the service would have had a favorable BCR in a year with more typical rainfall. For this exercise, in Panel C, we recalculate the BCR using the estimated benefit for excess rainfall-affected farmers from the 2021 experiment and the corresponding cost figures, but with the assumption that the share of farmers affected by the shock matched the 10-year average of the prevalence of excess rainfall in the state — 6%, rather than the 21% actually observed in 2021. Corresponding BCRs attenuate significantly.

In Panel D, we present the average BCRs for the three-year period from 2021 to 2023 using the average farmer reach recorded on a quarterly basis, the actual prevalence of excess rainfall over those three years, and the program cost over those three years. We continue to take the estimated effect on shock-affected farmers from the 2021 experiment. The scale of the service more than doubled, while the average program cost remained roughly the same. Even at the lower prevalence of weather shocks of $< 10\%$, the estimated BCRs are comparable to those in Panels A and B, driven by the increased coverage and falling annual costs.

Finally, Panel E attempts to capture the BCR of the program operated at scale. We use the treatment impact as estimated from the 2021 experiment and the average prevalence of shocks over the ten-year period 2014-2023. For this exercise, we assume that the farmer reach recorded at the end of 2023 and the program cost in 2023 reflect the long-term scale and cost of Ama Krushi based on its user growth-trajectory and spending pattern. While the farmer reach grew sharply from year to year through 2023, it has remained stable from 2023 to 2024, suggesting that the scale is approaching its steady state. The total cost also remained constant from 2023 to 2024. With the service reaching near 7 million farmers at an annual budget of US$ 1 million, the BCR reaches 13:1 in the long term.

The five scenarios presented in Table 13 illustrate some of the drivers of the BCR for this service. Because the largest benefits accrue to farmers affected by excess rainfall shocks, the extent of shocks largely determines the benefits of the service. The first year of the experiment, 2021, was an outlier in that 21% of farmers were affected by excess rainfall compared to 6% in an average year. This means that looking only at the BCR for 2021 could overstate the value of the service. However, Ama Krushi had not yet reached its steady state scale as of 2021. Coverage increased and costs fell rapidly over the next two years, so the 2021 BCR reflected higher costs than what will characterize the service's long-run operations. Taking into account the average incidence of excess rainfall and the expected coverage and costs of the program in the long run suggests an expected BCR of about 13. Using the preferred (winsorized) measure of farmers' costs to compute profits, the 95% CI for the BCR excludes values lower than 3.3. Therefore, we expect that this program provides a good value-proposition in the long run even though it primarily benefits farmers affected by relatively scarce excess rainfall shocks.

One limitation of this exercise is that our estimates of the impact on profits come from our evaluation data from five out of thirty districts in the one agricultural season for which we have experimental treatment effects and data on production costs. While we implemented the experiment with a second cohort in the following year, we did not collect data on production costs in the second year. The nature of the shocks that affected farmers in the sample were different in the second year, so while we find that Ama Krushi increased yields and harvests for farmers in areas affected by inadequate rainfall in 2022, we lack the data to calculate effects on profits and therefore to incorporate the benefits to drought-affected farmers (or those affected by other shocks that were not realized in 2021, the only year in which we collected all of the data necessary to measure profits) into our long-run BCR calculations.

8 Conclusion

This study evaluates the impact of a digital advisory service reaching millions of smallholder farmers in Odisha, India. Existing evidence demonstrates the effectiveness of digital extension services in promoting the adoption of recommended practices in the smallholder farmer context. However, few studies measure the impact of a service that leverages localized, real-time information to customize advice on a wide range of agricultural topics at scale.

Our findings offer initial evidence that such a service can increase farmers' capability to mitigate adverse effects of *some* production risks. Our analysis shows that the digital advisory service reduced the likelihood of severe crop loss and that the positive impact of the service is concentrated in areas hit by weather shocks, including excess rainfall in one year and scarce rainfall in another year. In contrast, we do not find a significant impact for farmers in areas that were unaffected by shocks during the evaluation period, despite the similar level of service usage. These results suggest that the impact of and who benefits from the service can vary from season to season, depending on the presence of weather shocks.

There are several plausible mechanisms through which the advisory helped farmers increase their resilience against weather shocks. The Ama Krushi advisory included precautionary advice on practices that enhance plant resilience to adverse events, including stress-tolerant seeds and micronutrient fertilizers; real-time prevention alerts for actions farmers could take to protect crops from severe damage immediately before adverse weather realization; and mitigation advice with damage-control measures immediately following adverse weather events, to minimize loss. While we find that the service overall improves farming knowledge and practices, our data on agricultural practices lack sufficient details to distinguish different mechanisms of impact. One productive area of future research is to better understand the heterogeneity of impact mechanisms by exploring the specific constraints farmers face. This could further improve the targeting of advisory content.

We also explore the use of remote sensing data as an alternative, more objective — and potentially more scalable — measure of the impact of the service. Our treatment impacts on the VI-predicted yield are positive in the first year, corroborating the survey results, and negative and insignificant in the second year, a deviation from the survey results. Our analysis suggests that the inconsistent, second-year results may be driven by farmers changing cultivation areas across seasons, consequently resulting in an inaccurate prediction of yield when relying on baseline plot boundary data. While these methodological challenges pose barriers to using satellites to measure program impact in a randomized evaluation, we demonstrate that they can offer valuable insights on differential attrition and survey sample selection to assess the validity of survey results.

There are untapped opportunities to improve smallholder farmers' resilience to shocks by making more targeted information available. Our calculations suggest that the long-run benefits generated by Ama Krushi would be an order of magnitude larger than the cost of the at-scale service operation in 2023 with near seven million farmers in the state. Further work to uncover heterogeneous mechanisms of impact across different types of information could inform future program designs, as advanced technologies expand the scope of localized information that can be made available in real time.

9 Figures

Figure 1: The timeline of the study

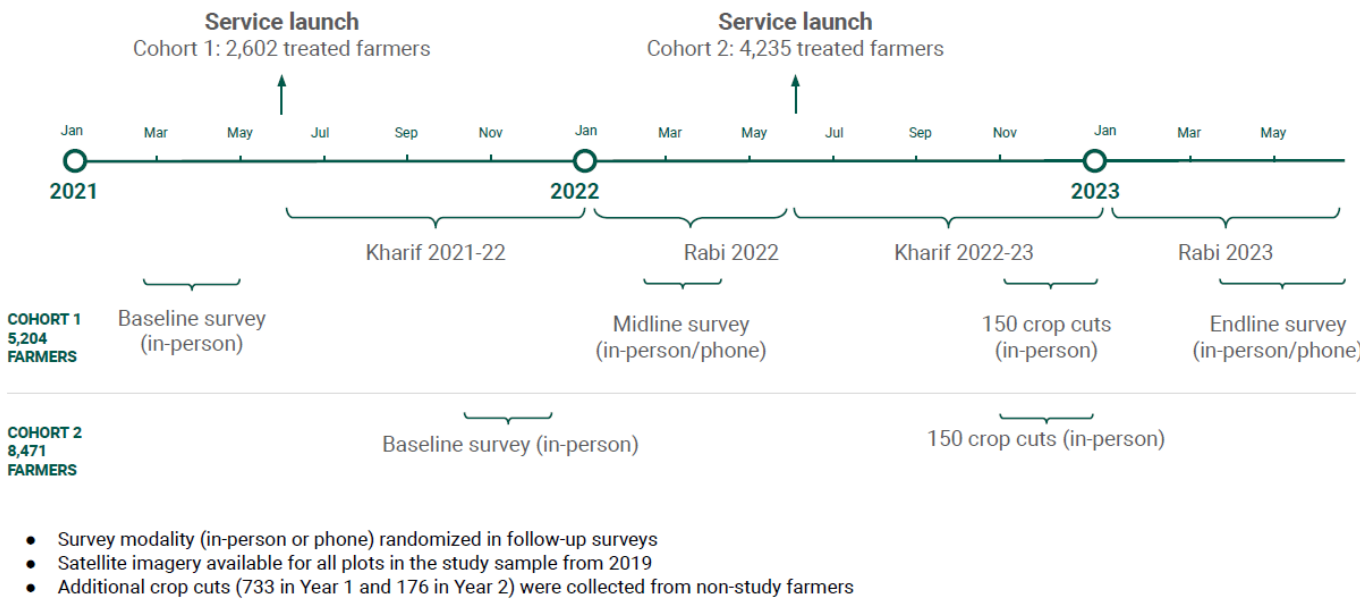
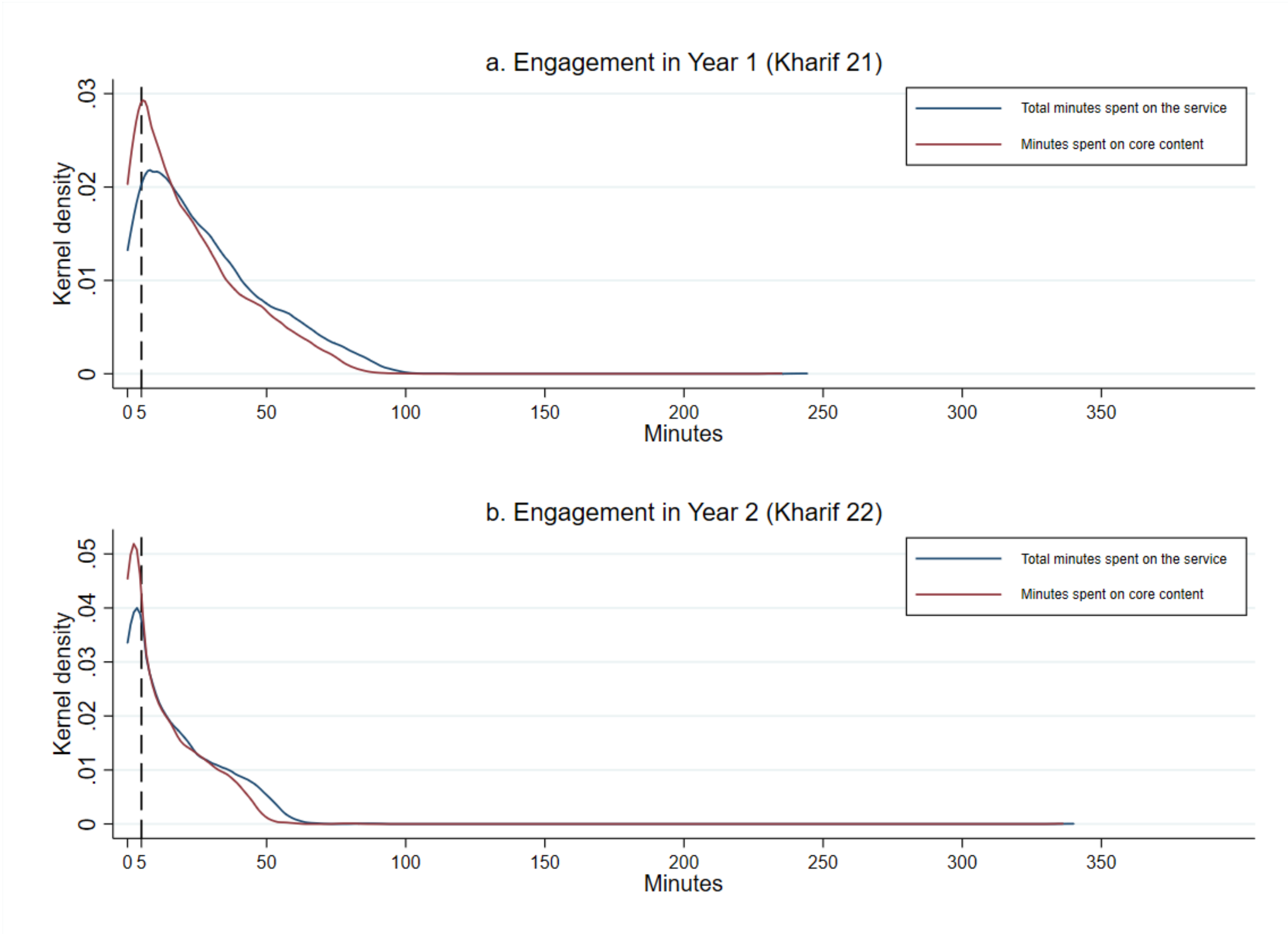


Figure 2: The distribution of Cohort 1 farmers' engagement with the digital extension service



Note: Data are from Cohort 1 farmers. The figures present the distributions of time spent on the digital agricultural extension service. Total minutes include the time spent on listening to all push calls and the time spent on the hotline. Minutes spent on core content include the time spent on listening to farming advice (as opposed to notices and information on government schemes) and the hotline.

Figure 3: The distribution of self-reported total rice yield (kg/hectare)

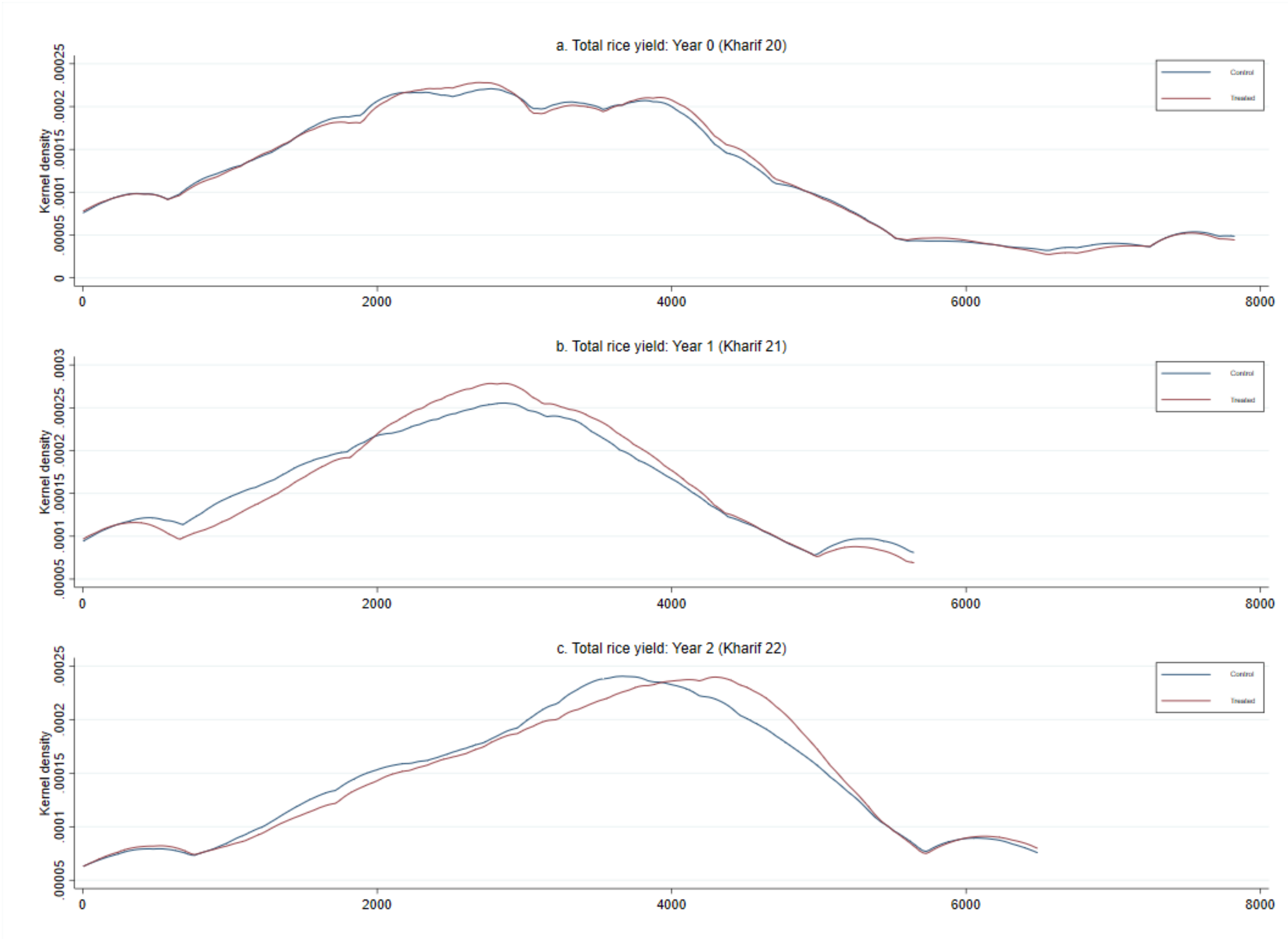
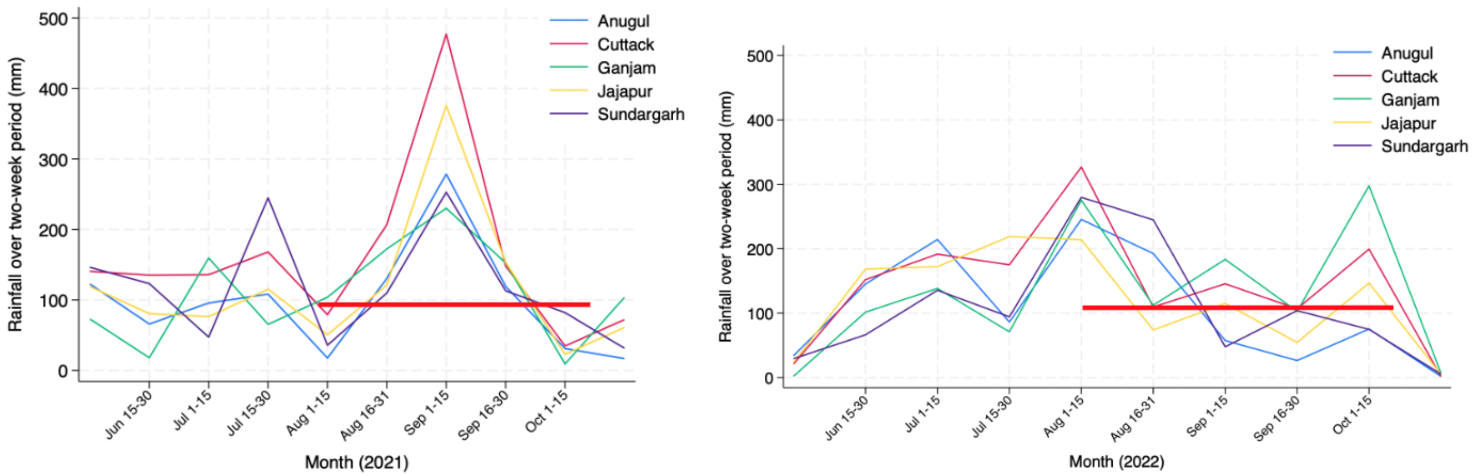


Figure 4: Rainfall patterns during the Kharif season in 2021 and 2022



Note: The figures show the mean block-level rainfall during the course of the Kharif season by district. The horizontal red line indicates the level of rainfall needed for rice cultivation. In 2021, the spike in the time window at September 1-15 indicates the extremely heavy rainfall that submerged a large proportion of rice fields in some study blocks. In 2022, the mean rainfall between August 15 and October 15 dips below the indicative minimum required rainfall in three out of five districts. Source data: Odisha Daily Rainfall Monitoring System.

10 Tables

Table 1: Baseline statistics and the randomization balance

	Cohort 1			Cohort 2		
	(1) Control mean (SD)	(2) Treatment coefficient (SE)	(3) Total obs.	(4) Control mean (SD)	(5) Treatment coefficient (SE)	(6) Total obs.
Panel A: Baseline outcomes						
Primary rice plot yield	3,478	-95.56*	3,563	3,713	-47.04	7,386
	(2,207)	(55.99)		(2,146)	(37.08)	
Primary rice plot harvest	434.4	-2.407	5,087	457.8	2.940	7,855
	(387.7)	(9.165)		(405.7)	(6.325)	
Total rice cultivation area	0.786	-0.005	5,190	0.842	-0.012	8,428
	(0.537)	(0.015)		(0.616)	(0.012)	
Total rice yield	2,952	-19.28	4,751	3,246	-25.47	8,329
	(1,830)	(41.88)		(1,972)	(36.65)	
Total rice harvest	2,149	-32.76	4,762	2,387	-18.25	8,362
	(1,832)	(49.62)		(2,005)	(39.90)	
Value of harvest	29,424	-351.1	4,762	32,934	-189.3	8,362
	(25,864)	(666.4)		(28,543)	(555.8)	
Variable cost	19,765	144.8	5,200	24,819	16.86	8,438
	(15,110)	(407.3)		(18,425)	(352.9)	
Imputed profit	10,165	-505.0	4,759	8,366	-150.7	8,346
	(21,533)	(600.6)		(22,893)	(439.4)	
Panel B: Baseline controls						
Max. NDVI in Kharif 19	0.769	-0.000	5,204	0.761	-0.000	8,471
	(0.067)	(0.002)		(0.068)	(0.001)	
Max. NDVI in Kharif 20	0.781	-0.001	5,204	0.778	0.000	8,471
	(0.078)	(0.001)		(0.076)	(0.001)	
Female	0.170	0.012	5,204	0.131	-0.006	8,471
	(0.376)	(0.010)		(0.337)	(0.007)	
Age	43.94	0.279	5,204	44.13	0.169	8,471
	(9.800)	(0.267)		(10.07)	(0.209)	
Literacy index	-0.000	0.005	5,204	-0.000	0.003	8,471
	(1)	(0.027)		(1)	(0.022)	
Primary phone: solely own	0.390	-0.006	5,194	0.405	0.008	8,427
	(0.488)	(0.013)		(0.491)	(0.010)	
Primary phone: feature	0.674	-0.003	5,193	0.590	-0.006	8,427
	(0.469)	(0.012)		(0.492)	(0.011)	
HH wealth index	-0.000	0.001	5,204	0.000	-0.017	8,471
	(1)	(0.026)		(1)	(0.020)	
Access to irrigation	0.485	-0.006	5,201	0.466	-0.009	8,470
	(0.500)	(0.010)		(0.499)	(0.009)	
Compensated before service	0.496	-0.025*	5,204			
	(0.500)	(0.013)				
p-value of joint F-test		0.438			0.955	

Note: Data are from the baseline survey of Cohort 1 and Cohort 2 farmers. Baseline outcomes, except the imputed profit measure, are winsorized at the 95th percentile; imputed profit is winsorized at the 2.5th and the 97.5th percentiles. All Cohort 2 farmers were compensated before the service started. Fixed effects at the randomization strata level are included. Robust standard errors in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. The last row show the *p*-value of a joint test of all individual tests in the preceding rows.

Table 2: Farmers' engagement with the digital extension service in Year 1 (Kharif 2021)

	All treated farmers		Consent to midline		Difference
	(1)	(2)	(3)	(4)	(5)
	Mean	SD	Mean	SD	T-test
Panel A: Outbound push calls					
# of agricultural advisory calls sent	49.48	10.00	49.35	9.768	-0.130
# of push calls sent	52.46	10.11	52.33	9.875	-0.128
Avg. length (min) of push calls sent	1.817	0.136	1.816	0.141	-0.001
Total length (min) of push calls sent	95.89	18.83	95.67	18.40	-0.226
Registered for the service initially	0.995	0.073	0.994	0.076	-0.000
Listened >=1 ag. advisory call more than 10s	0.940	0.237	0.954	0.209	0.014**
Share of push calls picked up	0.546	0.280	0.574	0.271	0.028***
Avg. (cond.) listening rate of push calls	0.520	0.256	0.532	0.254	0.012
Share of push calls listend 80% content	0.243	0.224	0.263	0.227	0.019***
Total minutes spent on push calls	26.90	21.22	28.98	21.60	2.084***
Panel B: Inbound services					
Called inbound service	0.171	0.377	0.182	0.386	0.010
Accessed inbound features	0.042	0.201	0.047	0.211	0.004
Submitted questions to inbound	0.041	0.199	0.047	0.211	0.006
Received answers from inbound	0.038	0.191	0.043	0.203	0.005
Total minutes spent on inbound calls	0.830	5.162	0.912	5.579	0.081
Panel C: Total engagement					
Total minutes spent on the digital service	27.73	22.22	29.89	22.68	2.165***
Number of observations	2,602		2,078		4,680

Note: These statistics are generated from PxD's administrative data on service usage during the Kharif 2021 season among Cohort 1 farmers. The sample includes all Cohort 1 treated farmers in Columns (1) and (2) and all Cohort 1 treated farmers who consented to the midline survey in Columns (3) and (4). Column (5) reports the mean different between (1) and (3). Fourteen farmers opted out of the service at the time of the registration. Engagement metrics for these farmers take the value of zero.

Table 3: Farmers' engagement with the digital extension service in Year 2 (Kharif 2022)

	Cohort 1 farmers				Cohort 2 farmers		Differences	
	All treated farmers		Consent to endline		All treated farmers		(3) - (1)	(5) - (1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean	SD	Mean	SD	Mean	SD	T-test	T-test
Panel A: Outbound push calls								
# of agricultural advisory calls sent	46.21	6.349	45.96	6.374	45.17	6.585	-0.249	-1.037***
# of push calls sent	50.79	6.860	50.49	6.893	50.11	7.020	-0.297	-0.680***
Avg. length (min) of push calls sent	1.365	0.151	1.364	0.153	1.380	0.097	-0.000	0.015***
Total length (min) of push calls sent	70.17	9.645	69.76	9.689	69.47	9.96	-0.409	-0.699***
Registered for the service initially	0.988	0.109	0.988	0.110	0.995	0.069	-0.000	0.007***
Listened >=1 ag. advisory call more than 10s	0.835	0.371	0.849	0.358	0.928	0.258	0.014	0.093***
Share of calls picked up	0.398	0.287	0.417	0.287	0.494	0.264	0.019**	0.096***
Avg. (cond.) listening rate of calls	0.559	0.284	0.567	0.282	0.628	0.253	0.008	0.069***
Share of calls listened 80% content	0.207	0.229	0.222	0.235	0.282	0.235	0.015**	0.075***
Total minutes spent on push calls	15.81	15.07	16.79	15.35	21.38	15.26	0.982**	5.573***
Panel B: Inbound services								
Called inbound service	0.140	0.347	0.144	0.351	0.238	0.426	0.004	0.098***
Accessed inbound features	0.023	0.150	0.025	0.155	0.065	0.246	0.002	0.042***
Submitted questions to inbound	0.010	0.101	0.009	0.097	0.026	0.159	-0.001	0.016***
Received answers from inbound	0.007	0.081	0.006	0.075	0.026	0.158	-0.001	0.019***
Total minutes spent on inbound calls	0.533	6.738	0.424	2.418	0.920	3.610	-0.109	0.387***
Panel C: Total engagement								
Total minutes spent on the digital service	16.34	16.80	17.22	15.84	22.30	16.11	0.874*	5.960***
Number of observations	2,602		2,116		4,235		4,718	6,837

Note: This table reports the descriptive statistics from PxD's administrative data on service usage during the Kharif 2022 season among Cohort 1 and Cohort 2 farmers. The sample includes all Cohort 1 treated farmers in Columns (1)-(2), all Cohort 1 treated farmers who consented to the endline survey in Columns (3)-(4), and all Cohort 2 treated farmers in Columns (5)-(6). Column (7) reports the mean difference between (1) and (3); Columns (8) reports the mean difference between (1) and (5). Thirty one farmers from Cohort 1 and 20 farmers from Cohort 2 farmers opted out of the service before the start of the second season. Engagement metrics for these farmers take the value of zero.

Table 4: The impact of treatment on knowledge and adoption indices

	Midline knowledge		Midline adoption		Endline adoption	
	(1) Core index	(2) Transplanting index	(3) Core index	(4) Transplanting index	(5) Core index	(6) Transplanting index
Treated	0.114***	0.113**	0.068**	0.098***	0.050	0.111***
	(0.043)	(0.046)	(0.028)	(0.030)	(0.039)	(0.042)
N	2068	1782	4080	3479	2911	2457
R^2	0.017	0.024	0.015	0.014	0.019	0.019

Note: Data are from the midline and endline survey of Cohort 1 farmers. Knowledge questions in the midline survey and adoption questions in the endline survey were only collected in the in-person survey. The "core" summary index includes practices that are relevant to all farmers, and the analysis sample for this variable is the full set of farmers who consented to the respective survey. The "transplanting" summary index includes practices that are relevant to farmers that practice transplanting and the analysis sample for this variable includes farmers who consented to the respective survey and reported transplanting in the Kharif 2020-2021 season in the baseline survey. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). Fixed effects at the randomization strata level are included. Robust standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table 5: The impact of treatment on total rice farming

	(1) Planted paddy	(2) Cultivation area (ha)	(3) Yield (kg/ha)	(4) Harvest (kg)
Panel A: Year 1 - Kharif 21				
Treated	0.003	0.011	41.821	76.487*
	(0.308)	(0.390)	(0.331)	(0.094)
	[0.573]	[0.581]	[0.573]	[0.176]
N	4134	3835	3835	3835
R^2	0.005	0.434	0.053	0.395
Control mean	0.986	0.886	2713.196	2374.902
% Change	0.35	1.24	1.54	3.22
Panel B: Year 2 - Kharif 22				
Treated	0.000	0.018	64.653	126.225**
	(0.957)	(0.162)	(0.123)	(0.028)
	[0.942]	[0.317]	[0.238]	[0.020]
N	4170	3733	3733	3733
R^2	0.010	0.386	0.053	0.354
Control mean	0.955	0.776	3394.544	2564.961
% Change	0.04	2.29	1.90	4.92
Panel C: Pooled Year 1 and Year 2				
Treated	0.002	0.015	52.938*	101.730**
	(0.582)	(0.168)	(0.097)	(0.010)
	[0.631]	[0.317]	[0.182]	[0.012]
N	8304	7568	7568	7568
R^2	0.006	0.408	0.049	0.369
Control mean	0.970	0.832	3047.236	2468.081
% Change	0.21	1.79	1.74	4.12

Note: Data are from the midline and endline survey of Cohort 1 farmers. The dependent variables are (1) whether farmers planted rice, (2) the total rice cultivation area, (3) total rice yield (kg/ha) calculated using self-reported harvest and self-reported land size, and (4) self-reported total rice harvest (kg). Area, yield and harvest outcomes are winsorized at the 95th percentile. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). For the total rice harvest amount outcome (Column 4), two additional variables – total rice farming area reported at the baseline and whether its value is imputed with block median – are controlled. Fixed effects at the randomization strata level are included. Robust standard errors are calculated in Panel A and B and standard errors are clustered at the farmer level in Panel C: * p<0.10, ** p<0.05, *** p<0.01. Their corresponding *p*-values are reported in parentheses, and Romano and Wolf's step-down adjusted *p*-values are reported in brackets. % changes represent the treatment impacts over the control means.

Table 6: The impact of treatment on rice revenue and profit (Year 1)

	Full sample	In-person sample				
	(1)	(2)	(3)	(4)	(5)	(6)
	Harvest (kg)	Harvest (kg)	Reported revenue (Rs)	Value of harvest (Rs)	Variable cost (Rs)	Imputed profit (Rs)
Treated	76.487*	107.991	656.696	1455.572	820.334	326.414
	(0.094)	(0.108)	(0.460)	(0.141)	(0.139)	(0.726)
		[0.383]	[0.856]	[0.413]	[0.487]	[0.856]
N	3835	1929	1929	1929	1929	1929
R^2	0.395	0.422	0.355	0.418	0.473	0.113
Control mean	2374.902	2238.218	16606.245	32774.124	25634.269	7706.750
% Change	3.22	4.82	3.95	4.44	3.20	4.24

Note: Data are from the midline survey of Cohort 1 farmers. The dependent variables are (1) and (2) total rice harvest amount (kg), (3) self-reported revenue from rice sales (INR), (4) total value of rice harvest calculated using median sales prices in farmers' location (INR), (5) total variable cost of rice cultivation (INR), summation of raw cost items, and (6) imputed profit of rice harvest (total value of harvest minus total variable cost, INR). Harvest, revenue, value of harvest, and cost outcomes are winsorized at the 95th percentile, and imputed profit outcome is winsorized at the 2.5th and the 97.5th percentiles. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). Fixed effects at the randomization strata level are included. Robust standard errors are calculated: * p<0.10, ** p<0.05, *** p<0.01. Their corresponding p-values are reported in parentheses, and Romano and Wolf's step-down adjusted p-values are reported in brackets. % changes represent the treatment impacts over the control means.

Table 7: The impact of treatment on rice loss (Year 2)

	Any rice crop loss					Severe rice crop loss (more than 50%)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All events	Flood	Other weather events	Pests & diseases	Animals	All events	Flood	Other weather events	Pests & diseases	Animals
Treated	-0.030**	-0.014	-0.030**	-0.007	0.007	-0.020*	0.003	-0.026***	-0.014**	-0.001
	(0.041)	(0.103)	(0.026)	(0.591)	(0.346)	(0.060)	(0.589)	(0.003)	(0.039)	(0.844)
	[0.030]	[0.094]	[0.014]	[0.848]	[0.541]	[0.046]	[0.848]	[0.002]	[0.030]	[0.848]
N	3973	3973	3973	3973	3973	3973	3973	3973	3973	3973
R^2	0.028	0.021	0.010	0.017	0.004	0.025	0.020	0.013	0.020	0.005
Control mean	0.612	0.161	0.317	0.199	0.074	0.208	0.091	0.105	0.053	0.021
% Change	-4.86	-8.51	-9.45	-3.52	9.96	-9.55	3.66	-24.30	-25.98	-4.36

Note: Data are from the endline survey of Cohort 1 farmers. The dependent variables are (1) a dummy variable indicating whether farmers experienced any rice loss during the 2022-2023 Kharif season, (6) a dummy variable indicating whether farmers experienced severe rice loss that was more than 50% of crops, and (2)-(5) and (7)-(10) dummy variables indicating rice lost due to specific reasons. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). Fixed effects at the randomization strata level are included. Robust standard errors are calculated: * p<0.10, ** p<0.05, *** p<0.01. Their corresponding p-values are reported in parentheses, and Romano and Wolf's step-down adjusted p-values are reported in brackets. % changes represent the treatment impacts over the control means.

Table 8: The heterogeneous treatment impact on total rice farming by weather shocks (Year 1)

	(1) Yield (kg/ha)	(2) Harvest (kg)	(3) Imputed profit (Rs)
Panel A: Excess rainfall			
Treated	-7.282	28.023	-352.758
	(0.879)	(0.450)	(0.639)
	[1.000]	[0.994]	[0.994]
Treatment × Excess rainfall	172.790**	170.526**	2592.727*
	(0.048)	(0.018)	(0.075)
	[0.395]	[0.204]	[0.529]
N	3835	3835	1929
Ctrl. mean (No excess rainfall)	2826.766	2474.509	10323.159
Ctrl. mean (Excess rainfall)	2424.854	2122.009	-9.439
Total effect for excess rainfall areas	165.508**	198.549***	2239.970*
	(0.022)	(0.001)	(0.071)
	[0.226]	[0.074]	[0.529]
Panel B: Cyclone			
Treated	0.491	39.169	-209.916
	(0.993)	(0.519)	(0.854)
	[1.000]	[0.994]	[1.000]
Treatment × Cyclone	75.822	68.451	1072.380
	(0.369)	(0.344)	(0.451)
	[0.990]	[0.988]	[0.994]
N	3835	3835	1929
Ctrl. mean (No cyclone)	2624.047	2413.321	12662.899
Ctrl. mean (Cyclone)	2788.968	2342.247	2707.877
Total effect for cyclone areas	76.313	107.620**	862.464
	(0.250)	(0.014)	(0.349)
	[0.948]	[0.196]	[0.990]

Note: Data are from the midline survey of Cohort 1 farmers. The dependent variables are (1) total rice yield (kg/ha), (2) total rice harvest (kg), and (3) imputed profit of rice harvest, in which the value of the harvest is calculated using median prices in the farmers' location (INR). Harvest and yield outcomes are winsorized at the 95th percentile, and the imputed profit outcome is winsorized at the 2.5th and the 97.5th percentiles. Weather shocks are defined as follows: "excess rainfall" refers to more than 250 mm rainfall over two days in the period of August 16 to October 15 in 2021 (i.e., the main growing stage) and is defined at the block level; "cyclone" refers to more than 50 mm rainfall over two days in the period of December 4 to December 10 in 2021 (i.e. the period of Cyclonic Storm Jawad) and is defined at the block level. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). For the total rice harvest and imputed profit outcomes, two additional variables - total rice farming area reported at the baseline and whether its value is imputed with block median - are controlled. Fixed effects at the randomization strata level are included. Standard errors are clustered at block level: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Their corresponding p-values are reported in parentheses, and Romano and Wolf's step-down adjusted p-values are reported in brackets.

Table 9: The heterogeneous treatment impact on total rice farming by weather shocks (Year 2)

	(1) Yield (kg/ha)	(2) Harvest (kg)	(3) Severe rice crop loss
Panel A: River flooding			
Treated	85.947**	156.609**	-0.026**
	(0.048)	(0.032)	(0.031)
	[0.489]	[0.377]	[0.377]
Treatment × River flooding	-104.654	-149.368	0.032
	(0.430)	(0.455)	(0.199)
	[0.994]	[0.994]	[0.918]
N	3733	3733	3973
Ctrl. mean (No river flooding)	3615.080	2720.652	0.152
Ctrl. mean (River flooding)	2519.552	1947.246	0.451
Total effect for river flooding areas	-18.707	7.241	0.006
	(0.882)	(0.969)	(0.785)
	[1.000]	[1.000]	[0.998]
Panel B: Scarce rainfall			
Treated	49.693	85.687	-0.006
	(0.311)	(0.279)	(0.595)
	[0.978]	[0.968]	[0.994]
Treatment × Scarce rainfall	57.248	155.125	-0.052**
	(0.574)	(0.422)	(0.016)
	[0.994]	[0.992]	[0.196]
N	3733	3733	3973
Ctrl. mean (No scarce rainfall)	3561.405	2737.797	0.183
Ctrl. mean (Scarce rainfall)	2925.115	2078.723	0.278
Total effect for scarce rainfall areas	106.942	240.812	-0.058***
	(0.250)	(0.179)	(0.002)
	[0.948]	[0.866]	[0.074]

Note: Data are from the endline survey of Cohort 1 farmers. The dependent variables are (1) total rice yield (kg/ha), (2) total rice harvest (kg), and (3) a dummy indicating whether the respondent experienced severe (>50%) rice loss. Harvest and yield outcomes are winsorized at the 95th percentile. Weather shocks are defined as follows: "river flooding" refers to river flooding in August 2022 and is defined at the panchayat level; and "scarce rainfall" refers to less than 450 mm rainfall over the period of August 16 to October 15 in 2022 (i.e., the main growing stage) and is defined at the block level. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). For the total rice harvest outcome, two additional variables – total rice farming area reported at the baseline and whether its value is imputed with block median – are controlled. Fixed effects at the randomization strata level are included. Standard errors are clustered at the panchayat level for the "river flooding" panel and at the block level for the "scarce rainfall" panel: * p<0.10, ** p<0.05, *** p<0.01. Their corresponding *p*-values are reported in parentheses, and Romano and Wolf's step-down adjusted *p*-values are reported in brackets.

Table 10: The heterogeneous treatment impact on total rice farming by baseline yield

	Midline					Endline				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Adoption core index	Cultivation area (ha)	Yield (kg/ha)	Harvest (kg)	Imputed profit (Rs)	Adoption core index	Cultivation area (ha)	Yield (kg/ha)	Harvest (kg)	Severe rice crop loss
Treated	0.046	0.036**	26.131	157.418**	1391.575	0.088*	0.036**	31.576	166.288**	-0.015
	(0.247)	(0.043)	(0.676)	(0.014)	(0.251)	(0.097)	(0.033)	(0.568)	(0.037)	(0.301)
	[0.824]	[0.104]	[0.986]	[0.018]	[0.824]	[0.333]	[0.066]	[0.974]	[0.080]	[0.888]
Treated × High productivity	0.048	-0.051**	32.448	-167.338*	-2219.132	-0.079	-0.038	68.165	-82.557	-0.010
	(0.393)	(0.044)	(0.705)	(0.061)	(0.231)	(0.307)	(0.124)	(0.416)	(0.466)	(0.638)
	[0.950]	[0.104]	[0.986]	[0.164]	[0.796]	[0.888]	[0.439]	[0.950]	[0.952]	[0.986]
N	4080	3835	3835	3835	1929	2911	3733	3733	3733	3973
R^2	0.016	0.435	0.054	0.395	0.114	0.020	0.386	0.053	0.354	0.025
Ctrl mean (Low productivity)	-0.011	0.857	2569.700	2168.568	5548.136	-0.042	0.746	3231.132	2337.979	0.247
Ctrl mean (High productivity)	0.012	0.918	2874.389	2606.683	10120.473	0.046	0.808	3573.607	2813.683	0.165
Total effect for high productivity	0.093**	-0.015	58.578	-9.920	-827.557	0.009	-0.002	99.742	83.731	-0.025*
	(0.020)	(0.399)	(0.319)	(0.876)	(0.557)	(0.876)	(0.912)	(0.113)	(0.303)	(0.099)
	[0.036]	[0.950]	[0.888]	[0.988]	[0.974]	[0.988]	[0.988]	[0.385]	[0.888]	[0.341]

Note: Data are from the midline and endline survey of Cohort 1 farmers. The dependent variables are (1) and (6) summary index of adoptions (core practices), (2) and (7) total rice cultivation area (ha), (3) and (8) total rice yield (kg/ha), (4) and (9) total rice harvest (kg), (5) imputed profit of rice harvest, in which the value of the harvest is calculated using median prices in farmers' location (INR), and (10) a dummy indicating whether the respondent experienced severe (>50%) rice loss. Area, yield and harvest outcomes are winsorized at the 95th percentile, and imputed profit outcome is winsorized at the 2.5th and the 97.5th percentiles. "High yield" refers to the respondent's baseline rice yield of the Kharif 2020-2021 season being above the median value. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). For the harvest and profit outcomes, two additional variables – the total rice farming area reported at the baseline and whether its value is imputed with the block median – are controlled. Fixed effects at the randomization strata level are included. Robust standard errors are calculated: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Their corresponding p-values are reported in parentheses, and Romano and Wolf's step-down adjusted p-values are reported in brackets.

Table 11: The impact of treatment on primary rice plot farming

	(1) Planted rice	(2) Cultivation area (ha)	(3) Harvest (kg)	(4) Reported yield (reported area)	(5) Reported yield (measured area)	(6) VI-predicted yield (measured area)
Panel A: Year 1 - Kharif 21						
Treated	-0.001	0.000	11.423	103.289	104.702	25.636**
	(0.007)	(0.003)	(12.168)	(64.065)	(112.178)	(10.430)
N	4008	3220	3220	3220	3220	3220
R^2	0.007	0.441	0.354	0.048	0.048	0.119
Control mean	0.943	0.164	473.989	2936.003	4352.503	2998.045
% Change	-0.16	0.24	2.41	3.52	2.41	0.86
Panel B: Year 2 - Kharif 22						
Treated	-0.003	-0.002	7.533	68.160	-8.639	-12.616
	(0.008)	(0.002)	(11.279)	(60.740)	(108.540)	(10.796)
N	3935	3194	3194	3194	3194	3194
R^2	0.009	0.493	0.415	0.043	0.050	0.129
Control mean	0.942	0.152	532.180	3682.927	5053.424	4036.924
% Change	-0.35	-1.48	1.42	1.85	-0.17	-0.31
Panel C: Pooled Year 1 and Year 2						
Treated	-0.002	-0.001	9.283	85.457*	48.906	6.906
	(0.005)	(0.002)	(9.158)	(43.965)	(84.772)	(7.467)
N	7943	6414	6414	6414	6414	6414
R^2	0.006	0.465	0.380	0.042	0.045	0.120
Control mean	0.943	0.158	502.623	3303.537	4697.401	3509.239
% Change	-0.24	-0.54	1.85	2.59	1.04	0.20

Note: Data are from the midline and endline survey of Cohort 1 farmers. The dependent variables all refer to the primary rice plot and are (1) whether farmers planted rice, (2) the total rice cultivation area (ha), (3) self-reported rice harvest (kg), (4) rice yield (kg/ha) calculated using self-reported harvest and self-reported land size, (5) rice yield (kg/ha) calculated using self-reported harvest and GPS measured land size, and (6) VI-predicted yield (kg/ha). Area, yield and harvest outcomes are winsorized at the 95th percentile. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). For the primary rice-plot harvest outcome, one additional variable – primary rice-plot land size measured by GPS at the baseline – is controlled. Fixed effects at the randomization strata level are included. Standard errors are robust in Panel A and B and clustered at the farmer level in Panel C, reported in parentheses: * p<0.10, ** p<0.05, *** p<0.01. % changes represent the treatment impacts over the control means.

Table 12: The impact of treatment on the primary rice-plot yield (VI-predicted yield)

	Analysis sample			Full sample		
	(1) In-person sample (profit/ adoption)	(2) Total rice sample (yield)	(3) Survey consent sample	(4) Cohort 1 sample	(5) Cohort 2 sample	(6) Full RCT sample
Panel A: Year 1 - Kharif 21						
Treated	11.476	17.693*	19.422**	13.787*		
	(11.648)	(9.146)	(8.770)	(7.624)		
N	1929	3835	4140	5204		
R^2	0.124	0.107	0.106	0.100		
Control mean	3005.853	2991.779	2978.179	2962.892		
Panel B: Year 2 - Kharif 22						
Treated	-12.089	-11.216	-11.921	-12.899	-5.470	-8.398
	(12.387)	(9.450)	(9.794)	(8.588)	(6.077)	(5.164)
N	2724	3733	4194	5204	8463	13667
R^2	0.133	0.130	0.121	0.115	0.115	0.114
Control mean	3985.455	3984.432	3940.554	3905.160	3826.659	3856.543

Note: Data are from the midline and endline survey of Cohort 1 and 2 farmers. Different columns use different samples. The dependent variables are VI-predicted yields. All regressions control for vegetation indices from the one pre-intervention Kharif seasons (2019) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention. Fixed effects at the randomization strata level are included. Bootstrap standard errors are reported in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table 13: Estimated Benefit-Cost Ratios (BCRs)

	(1) Farmer reach (in millions)	(2) Proportion affected by excess rainfall	(3) Total benefit (in US$ millions)	(4) Total cost (in US$ millions)	(5) Benefit-cost ratio central estimates (95% CI)
Estimated using the average impact on profit					
Panel A. BCR in 2021					
Profit measure:					
As specified in PAP	1.371	/	6.056	0.972	6.23 (-28.51, 40.97)
Using winsorized cost components	1.371	/	8.827	0.972	9.08 (-22.69, 40.86)
Unsold harvest valued at retail price	1.371	/	18.090	0.972	18.62 (-30.85, 68.08)
Reference year:	2021	/	2021	2021	
Estimated using the impact on profit in areas hit by excess rainfall					
Panel B. BCR in 2021					
Profit measure:					
As specified in PAP	1.371	21%	8.882	0.972	9.14 (0.02, 18.26)
Using winsorized cost components	1.371	21%	9.931	0.972	10.22 (3.26, 17.18)
Unsold harvest valued at retail price	1.371	21%	14.078	0.972	14.49 (-0.47, 29.44)
Reference year:	2021	2021	2021	2021	
Panel C. BCR in 2021, adjusted for the long-run prevalence of excess rainfall					
Profit measure:					
As specified in PAP	1.371	6%	2.662	0.972	2.74 (0.01, 5.47)
Using winsorized cost components	1.371	6%	2.976	0.972	3.06 (0.98, 5.15)
Unsold harvest valued at retail price	1.371	6%	4.219	0.972	4.34 (-0.14, 8.82)
Reference year:	2021	2014-2023	2021	2021	
Panel D. 3-year BCR, 2021-2023					
Profit measure:					
As specified in PAP	3.173	8%	7.394	0.998	7.41 (0.01, 14.81)
Using winsorized cost components	3.173	8%	8.267	0.998	8.29 (2.64, 13.93)
Unsold harvest valued at retail price	3.173	8%	11.720	0.998	11.75 (-0.38, 23.88)
Reference year:	2021-2023	2021-2023	2021-2023	2021-2023	
Panel E. Long-run BCR					
Profit measure:					
As specified in PAP	6.872	6%	11.939	1.016	11.75 (0.02,23.48)
Using winsorized cost components	6.872	6%	13.348	1.016	13.14 (4.19,22.09)
Unsold harvest valued at retail price	6.872	6%	18.923	1.016	18.62 (-0.60,37.85)
Reference year:	2023	2014-2023	2023	2023	

Note: The average impacts on profit are from Column (6) in Table 6 and Appendix Table A8; impacts on profit in excess rainfall areas are from Table 8 and Appendix Table A9. They are estimated impacts on profit per farmer in the Kharif 2021 season. "Farmer reach" refers to the number of farmers registered for the service in Odisha. "Proportion affected by excess rainfall" refers to the share of farmers affected by excess rainfall in Odisha. The long-run prevalence of excess rainfall is calculated using the past 10 years' (2014-2023) prevalence levels. The benefit-cost ratio is calculated as the total estimated benefit dividing by the total estimated cost. Numbers in the parentheses stand for the 95% confidence interval. The total benefit of the service equals the treatment impacts on profit per farmer in the Kharif 2021 season multiplied by the number of relevant farmers reached by the service in that year, which is calculated as the number of farmers registered for the service in that year for the average impact, and that number multiplied by the share of farmers in Odisha affected by the excess rainfall shock in that year for the impact in excess rainfall areas.

References

Abate, Gashaw T., Kibrom A. Abay, Jordan Chamberlin, Yumna Kassim, David J. Spielman, and Martin Paul Jr Tabe-Ojong. 2023. "Digital Tools and Agricultural Market Transformation in Africa: Why Are They Not at Scale Yet, and What Will It Take to Get There?" *Food Policy* 116:102439. https://doi.org/10.1016/j.foodpol.2023.102439.

Abay, Kibrom A. 2020. "Measurement Errors in Agricultural Data and Their Implications on Marginal Returns to Modern Agricultural Inputs." *Agricultural Economics* 51:323—341. https://doi.org/10.1111/agec.12557.

Abay, Kibrom A., Gashaw T. Abate, Christopher B. Barrett, and Tanguy Bernard. 2019. "Correlated Non-classical Measurement Errors, 'Second Best' Policy Inference, and the Inverse Size-productivity Relationship in Agriculture." *Journal of Development Economics* 139:171–184. https://doi.org/10.1016/j.jdeveco.2019.03.008.

Aker, Jenny C. 2011. "Dial "A" for Agriculture: A Review of Information and Communication Technologies for Agricultural Extension in Developing Countries." *Agricultural Economics* 42 (6): 631–647. https://doi.org/10.1111/j.1574-0862.2011.00545.x.

Aker, Jenny C., Ishita Ghosh, and Jenna Burrell. 2016. "The Promise (and Pitfalls) of ICT for Agriculture Initiatives." *Agricultural Economics* 47:35–48. https://doi.org/10.1111/agec.12301.

Anderson, Jock R., and Gershon Feder. 2004. "Agricultural Extension: Good Intentions and Hard Realities." *The World Bank Research Observer* 19:41—60. https://doi.org/10.1093/wbro/lkh013.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–1495. https://doi.org/10.1198/016214508000000841.

Arouna, Aminou, Jeffrey D. Michler, Wilfried G. Yergo, and Kazuki Saito. 2020. "One Size Fits All? Experimental Evidence on the Digital Delivery of Personalized Extension Advice in Nigeria." *American Journal of Agricultural Economics* 103:596–619. https://doi.org/10.1111/ajae.12151.

Ashraf, Nava, Xavier Giné, and Dean Karlan. 2009. "Finding Missing Markets (and a Disturbing Epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya." *American Journal of Agricultural Economics* 91:973–990. https://doi.org/10.1111/j.1467-8276.2009.01319.x.

Baul, Tushi, Dean Karlan, Kentaro Toyama, and Kathryn Vasilaky. 2024. "Improving Smallholder Agriculture via Video-based Group Extension." *Journal of Development Economics* 169:103267. https://doi.org/10.1016/j.jdeveco.2024.103267.

Bird, Samuel S., Michael R. Carter, Travis J. Lybbert, Mary Mathenge, Timothy Njagi, and Emilia Tjernström. 2022. "Filling a Niche? The Maize Productivity Impacts of Adaptive Breeding by a Local Seed Company in Kenya." *Journal of Development Economics* 157:102885. https://doi.org/10.1016/j.jdeveco.2022.102885.

Burke, Marshall, and David B. Lobell. 2017. "Satellite-based Assessment of Yield Variation and Its Determinants in Smallholder African Systems." *Proceedings of the National Academy of Sciences* 114:2189–2194. https://doi.org/10.1073/pnas.1616919114.

Burlig, Fiona, Amir Jina, Erin M. Kelley, Gregory Lane, and Harshil Sahai. 2024. *Long-Range Forecasts as Climate Adaptation: Experimental Evidence from Developing-Country Agriculture.* Working Paper 2024-21.

Campolo, Jake, Ivan Ortiz-Monasterio, David Guerena, and David Lobell. 2022. "Evaluating Maize Yield Response to Fertilizer and Soil in Mexico Using Ground and Satellite Approaches." *Field Crops Research* 276:108393. https://doi.org/10.1016/j.fcr.2021.108393.

Cole, Shawn, and Nilesh Fernando. 2021. "'Mobile'izing Agricultural Advice Technology Adoption Diffusion and Sustainability." *The Economic Journal* 131:192–219. https://doi.org/10.1093/ej/ueaa084.

Cole, Shawn, Tomoko Harigaya, Grady Killeen, and Aparn Krishna. 2025. "Using Satellites and Phones to Evaluate and Promote Agricultural Technology Adoption: Evidence from Smallholder Farms in India." *Journal of Development Economics* 176 (103463). https://doi.org/10.1016/j.jdeveco.2025.103463.

Consumer Affairs, Department of. n.d. *Retail Price: DCA: Daily: Rice: Odisha.* https://www.ceicdata.com/en/india/retail-price-department-of-consumer-affairs-agriculture-commodities-daily-by-states-rice/retail-price-dca-daily-rice-odisha. Accessed: July 2024. CEIC.

Das, S.R. 2012. *Rice in Odisha.* IRRI Technical Bulletins 287645. International Rice Research Institute. https://doi.org/10.22004/ag.econ.287645.

Deininger, Klaus, Calogero Carletto, Sara Savastano, and James Muwonge. 2012. "Can Diaries Help in Improving Agricultural Production Statistics? Evidence From Uganda." *Journal of Development Economics* 98 (1): 42–50. https://doi.org/10.1016/j.jdeveco.2011.05.007.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review* 101:2350–2390. https://doi.org/10.1257/aer.101.6.2350.

Dzanku, Fred Mawunyo, Robert Darko Osei, Paul Kwame Nkegbe, and Isaac Osei-Akoto. 2020. "Information Delivery Channels and Agricultural Technology Uptake: Experimental Evidence from Ghana." *European Review of Agricultural Economics* 49:82–120. https://doi.org/10.1093/erae/jbaa032.

Evenson, R. E., and D. Gollin. 2003. "Assessing the Impact of the Green Revolution, 1960 to 2000." *Science* 300:758–762. https://doi.org/10.1126/science.1078710.

Fabregas, Raissa, Tomoko Harigaya, Michael Kremer, and Ravindra Ramrattan. 2023. "Digital Agricultural Extension for Development." In *Introduction to Development Engineering,* edited by Temina Madon, Ashok J. Gadgil, Richard Anderson, Lorenzo Casaburi, Kenneth Lee, and Arman Rezaee, 187–219. Springer International Publishing. https://doi.org/10.1007/978-3-030-86065-3_8.

Fabregas, Raissa, Michael Kremer, Matthew Lowes, Robert On, and Giulia Zane. 2025. "Digital Information Provision and Behavior Change: Lessons from Six Experiments in East Africa." *American Economic Journal: Applied Economics* 17:527—66.

Fabregas, Raissa, Michael Kremer, and Frank Schilbach. 2019. "Realizing the Potential of Digital Development: The Case of Agricultural Advice." *Science* 366. https://doi.org/10.1126/science.aay3038.

Fafchamps, Marcel, and Bart Minten. 2012. "Impact of SMS-Based Agricultural Information on Indian Farmers." *The World Bank Economic Review* 26:383—414. https://doi.org/https://doi.org/10.1093/wber/lhr056.

Fu, Haishan, and Martien van Nieuwkoop. 2023. *Mind the Gap: Enabling Data-smart Agriculture for All.* https://blogs.worldbank.org/en/opendata/mind-gap-enabling-data-smart-agriculture-all. Accessed: 2024-07-10.

Gourlay, Sydney, Talip Kilic, and David B. Lobell. 2019. "A New Spin on an Old Debate: Errors in Farmer-reported Production and Their Implications for Inverse Scale – Productivity Relationship in Uganda." *Journal of Development Economics* 141:102376. https://doi.org/10.1016/j.jdeveco.2019.102376.

Guo, Zhe, Jordan Chamberlin, and Liangzhi You. 2023. "Smallholder Maize Yield Estimation Using Satellite Data and Machine Learning in Ethiopia." *Crop and Environment* 2:165–174. https://doi.org/10.1016/j.crope.2023.07.002.

Jain, Meha. 2020. "The Benefits and Pitfalls of Using Satellite Data for Causal Inference." *Review of Environmental Economics and Policy* 14:157–169. https://api.semanticscholar.org/CorpusID:214302516.

Jain, Meha, Balwinder-Singh, Preeti Rao, Amit K. Srivastava, Shishpal Poonia, Jennifer Blesh, George Azzari, Andrew J. McDonald, and David B. Lobell. 2019. "The Impact of Agricultural Interventions Can Be Doubled by Using Satellite Data." *Nature Sustainability* 2:931–934. https://doi.org/10.1038/s41893-019-0396-x.

Jain, Meha, Amit Srivastava, Balwinder-Singh, Rajiv Joon, Andrew McDonald, Keitasha Royal, Madeline Lisaius, and David Lobell. 2016. "Mapping Smallholder Wheat Yields and Sowing Dates Using Micro-Satellite Data." *Remote Sensing* 8:860. https://doi.org/10.3390/rs8100860.

Janvry, Alain de, Manaswini Rao, and Elisabeth Sadoulet. 2022. *Seeding the Seeds: Role of Social Structure in Agricultural Technology Diffusion.* Working Paper. https://manaswinirao.com/files/20220713_ss1diffusion.pdf.

Klompenburg, Thomas van, Ayalew Kassahun, and Cagatay Catal. 2020. "Crop Yield Prediction Using Machine Learning: A Systematic Literature Review." *Computers and Electronics in Agriculture* 177:105709. https://doi.org/10.1016/j.compag.2020.105709.

Koo, Jawoo. 2014. "Ex-ante Modeling Analysis of Maize Yield Responses to Nutrient Management Options in Malawi: Action Sites — Dedza & Ntcheu." In *Presented at the Africa RISING Monitoring and Evaluation Meeting.* Arusha: International Food Policy Research Institute.

Kubitza, Christoph, Vijesh V. Krishna, Urs Schulthess, and Meha Jain. 2020. "Estimating Adoption and Impacts of Agricultural Management Practices in Developing Countries Using Satellite Data. A Scoping Review." *Agronomy for Sustainable Development* 40. https://doi.org/10.1007/s13593-020-0610-2.

Lobell, David B, George Azzari, Marshall Burke, Sydney Gourlay, Zhenong Jin, Talip Kilic, and Siobhan Murray. 2019. "Eyes in the Sky, Boots on the Ground: Assessing Satellite- and Ground-Based Approaches to Crop Yield Measurement and Analysis." *American Journal of Agricultural Economics* 102:202–219. https://doi.org/10.1093/ajae/aaz051.

Mobarak, Ahmed Mushfiq, and Mark R Rosenzweig. 2013. "Informal Risk Sharing, Index Insurance, and Risk Taking in Developing Countries." *American Economic Review* 103:375–380. https://doi.org/10.1257/aer.103.3.375.

Nandi, Ravi, and Swamikannu Nedumaran. 2022. "Rural Market Food Diversity and Farm Production Diversity: Do They Complement or Substitute Each Other in Contributing to a Farm Household's Dietary Diversity?" *Frontiers in Sustainable Food Systems* 6. https://doi.org/10.3389/fsufs.2022.843697.

Okorie, Idika E., Emmanuel Afuecheta, and Saralees Nadarajah. 2023. "Time Series and Power Law Analysis of Crop Yield in Some East African Countries." Edited by Steven Arthur Loiselle. *PLOS ONE* 18:e0287011. https://doi.org/10.1371/journal.pone.0287011.

Planning and Convergence Department, Directorate of Economics and Statistics. 2021. *Odisha Economic Survey 2020-21.* Technical report. Bhubaneswar: Government of Odisha.

———. 2023. *Odisha Economic Survey 2022-23.* Technical report. Bhubaneswar: Government of Odisha.

Rice-based Cropping Systems. n.d. *Rice Farming in Odisha.* https://www.rkb-odisha.in/rice. Accessed: August 2024. International Rice Research Institutes (IRRI).

Romano, Joseph P., and Michael Wolf. 2016. "Efficient Computation of Adjusted p-Values for Resampling-Based Stepdown Multiple Testing." *Statistics & Probability Letters* 113:38–40. https://doi.org/10.1016/j.spl.2016.02.012.

Rosenzweig, Mark R, and Christopher Udry. 2019. "External Validity in a Stochastic World: Evidence from Low-Income Countries." *The Review of Economic Studies* 87:343–381. https://doi.org/10.1093/restud/rdz021.

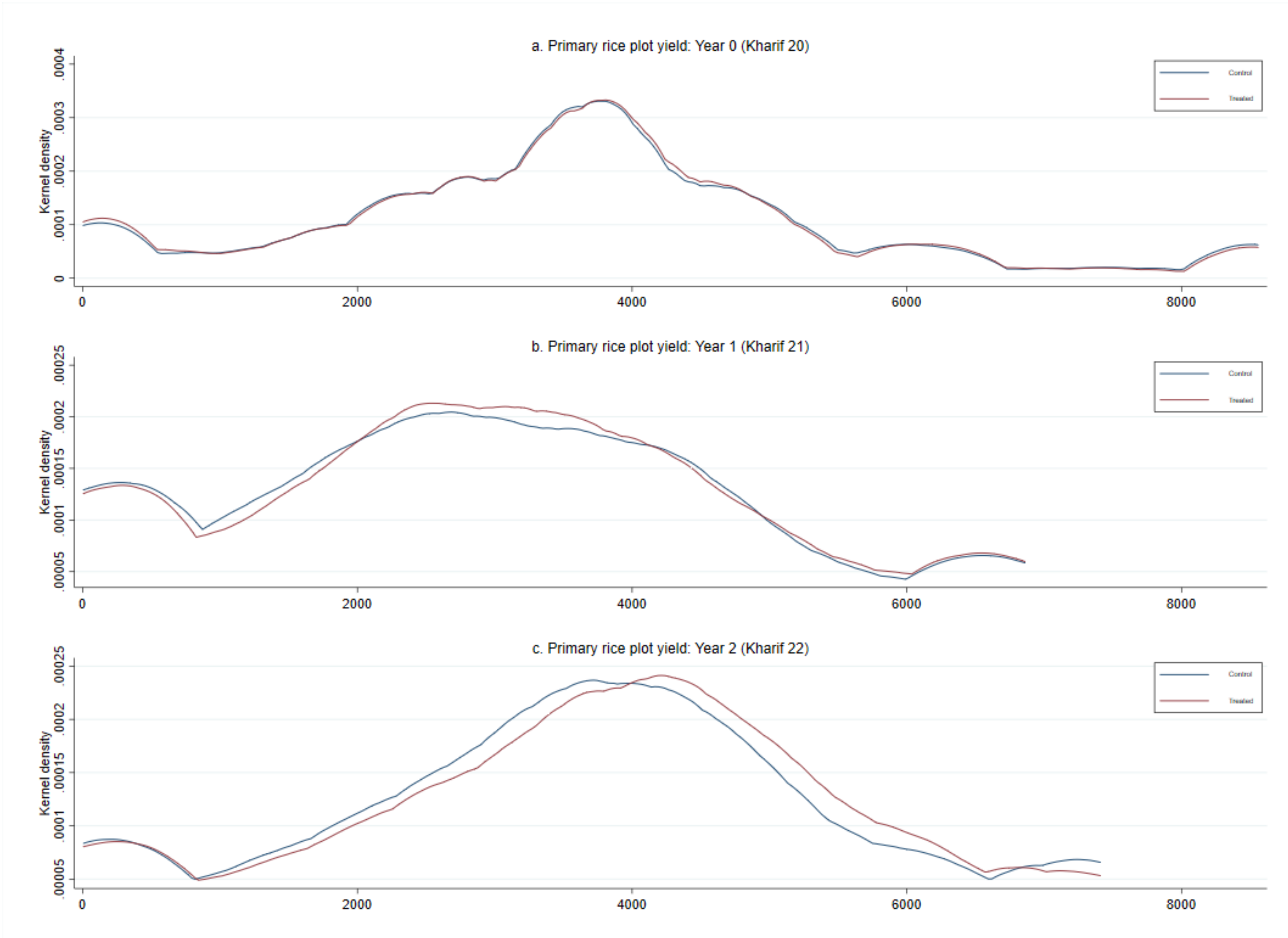
Spielman, David, Els Lecoutere, Simrin Makhija, and Bjorn Van Campenhout. 2021. "Information and Communications Technology (ICT) and Agricultural Extension in Developing Countries." *Annual Review of Resource Economics* 13:177–201. https://doi.org/10.1146/annurev-resource-101520-080657.

Yegbemey, Rosaine N., Gunther Bensch, and Colin Vance. 2023. "Weather Information and Agricultural Outcomes: Evidence From a Pilot Field Experiment in Benin." *World Development* 167. https://doi.org/10.1016/j.worlddev.2022.106178.

Appendix

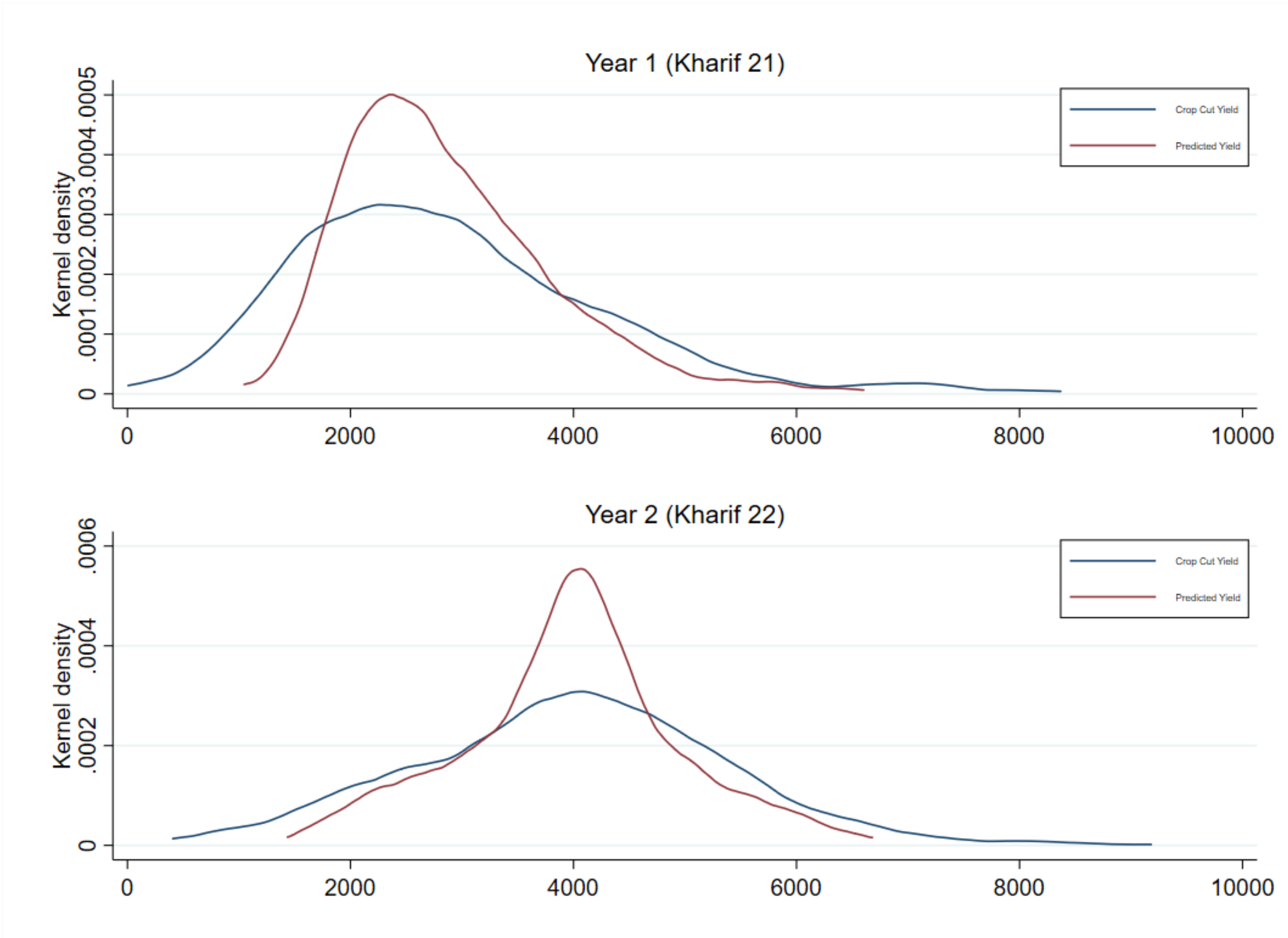
A Additional Figures and Tables

Figure A1: Distribution of self-reported primary rice plot yield (kg/hectare)



Note: Data are from the baseline, midline, and endline survey of Cohort 1 farmers. Yield outcomes are winsorized at the 95th percentile.

Figure A2: Distribution of crop-cut and predicted primary rice plot yield (CCE Sample)



Note: Data are from the crop cut exercises and yield prediction models.

Table A1: Attritions of follow-up surveys

	Consent to midline		Consent to endline		Consent to both	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.006	-0.161	0.014	-0.162	0.016	-0.066
	(0.009)	(0.138)	(0.012)	(0.152)	(0.012)	(0.163)
Treated × NDVI in Kharif 19		0.084		0.073		0.205
		(0.180)		(0.199)		(0.232)
Treated × NDVI in Kharif 20		0.164		0.137		-0.025
		(0.139)		(0.173)		(0.186)
Treated × Female		-0.003		0.002		-0.005
		(0.029)		(0.031)		(0.036)
Treated × Age		0.000		0.001		0.000
		(0.001)		(0.001)		(0.001)
Treated × Literacy index		0.010		-0.000		0.004
		(0.010)		(0.010)		(0.013)
Treated × Primary phone: solely own		-0.039		-0.024		-0.044
		(0.025)		(0.022)		(0.029)
Treated × Primary phone: feature		-0.005		0.024		-0.011
		(0.024)		(0.024)		(0.028)
Treated × HH wealth index		-0.014		0.009		-0.004
		(0.011)		(0.011)		(0.013)
Treated × Access to irrigation		0.003		-0.060**		-0.043
		(0.023)		(0.023)		(0.026)
Treated × Compensated before service		-0.023		-0.024		-0.033
		(0.021)		(0.022)		(0.025)
Max. NDVI in Kharif 19		0.114		-0.277*		-0.167
		(0.145)		(0.148)		(0.187)
Max. NDVI in Kharif 20		-0.246**		-0.269**		-0.221
		(0.119)		(0.133)		(0.152)
Female		0.017		0.062***		0.051*
		(0.021)		(0.022)		(0.026)
Age		0.001*		0.002***		0.003***
		(0.001)		(0.001)		(0.001)
Literacy index		-0.004		-0.001		-0.002
		(0.007)		(0.008)		(0.009)
Primary phone: solely own		0.022		0.015		0.032*
		(0.016)		(0.016)		(0.019)
Primary phone: feature		0.030*		0.000		0.045**
		(0.016)		(0.018)		(0.020)
HH wealth index		0.016**		0.005		0.017*
		(0.008)		(0.008)		(0.009)
Access to irrigation		0.006		0.053***		0.048**
		(0.017)		(0.018)		(0.020)
Compensated before service		0.115***		0.061***		0.124***
		(0.017)		(0.016)		(0.020)
N	5204	5204	5204	5204	5204	5204
R^2	0.000	0.024	0.000	0.018	0.000	0.023
Control mean	0.792	0.792	0.799	0.799	0.663	0.663
Treated × covariates	N	Y	N	Y	N	Y
p-value of joint F-test		0.495		0.228		0.350

Note: Data are from the midline and endline survey of Cohort 1 farmers. The dependent variables are whether farmers consented to (1)-(2) the midline surveys, (3)-(4) the endline surveys, and (5)-(6) both surveys. Fixed effects at the randomization strata level are included. Robust standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. The last row shows the p-value for a joint test of all interaction variables and the treatment indicator.

Table A2: Differential attrition:
Correlation between VI-predicted yield (in thousands) and survey responses

	(1) Responded to profit/ adoption questions	(2) Responded to profit/ adoption questions	(3) Responded to yield questions	(4) Consented to surveys
Panel A: Midline survey				
Treated	-0.024	-0.007	-0.033	-0.047
	(0.111)	(0.059)	(0.059)	(0.056)
Treated × VI-predicted yield	0.017	0.011	0.012	0.018
	(0.036)	(0.020)	(0.019)	(0.018)
VI-predicted yield	-0.030	0.005	0.010	0.004
	(0.031)	(0.019)	(0.020)	(0.018)
N	1885	5204	5204	5204
R^2	0.002	0.001	0.000	0.000
Avg response rate	0.823	0.371	0.737	0.796
Estimated at avg yield: control	0.811	0.358	0.735	0.793
Estimated at avg yield: treated	0.836	0.383	0.739	0.798
p-value: C=T	0.146	0.004	0.681	0.501
Panel B: Endline survey				
Treated	0.113	0.063	0.094	-0.005
	(0.098)	(0.079)	(0.072)	(0.075)
Treated × VI-predicted yield	-0.020	-0.011	-0.019	0.005
	(0.025)	(0.020)	(0.018)	(0.018)
VI-predicted yield	0.048	-0.008	0.029	-0.008
	(0.030)	(0.024)	(0.022)	(0.019)
N	3294	5204	5204	5204
R^2	0.002	0.001	0.001	0.000
Avg response rate	0.691	0.523	0.717	0.806
Estimated at avg yield: control	0.674	0.513	0.707	0.799
Estimated at avg yield: treated	0.707	0.533	0.728	0.813
p-value: C=T	0.040	0.116	0.112	0.235

Note: The dependent variables are indicators whether farmers responded to specific questions in the follow-up surveys, corresponding to various analysis samples we used. Column (1) restricts the sample to farmers who were assigned to in-person surveys initially, and Columns (2)-(4) use the full Cohort 1 sample. Fixed effects at the randomization strata level are included. Bootstrap standard errors are reported in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table A3: List of agricultural practices

Outcome	Definition
Suitable seed	Adopted suitable seeds that have recommended maturity duration: seeds that mature in 100-120 days for upland areas, in 120-140 days for medium land areas, and over 140 days for lowland areas.
Seed replacement	Adopted correct seed replacement rate: local seeds may be reused multiple times, variety seeds up to three times, and hybrid seeds only once.
Seed treatment	Treated seeds with chemicals.
Pesticide	Adopted any pesticide.
Herbicide	Adopted any recommended herbicide.
Nursery fertilizer fym	Adopted the recommended fertilizer in nursery: FYM.
Nursery fertilizer dap	Adopted the recommended fertilizer in nursery: DAP.
Nursery fertilizer mop	Adopted the recommended fertilizer in nursery: MOP.
Nursery fertilizer urea	Adopted the recommended fertilizer in nursery: Urea.
Fertilizer dap	Applied the recommended fertilizer: DAP.
Fertilizer mop	Applied the recommended fertilizer: MOP.
Fertilizer urea	Applied the recommended fertilizer: Urea.
Micronutrient zinc	Applied the recommended micronutrient fertilizer: Zinc.
Line method	Adopted line method for transplanting or broadcasting.
Transplanting time	Transplanted seedling at the recommended time (i.e., between July 1 and August 15).

Table A4: The impact of treatment on adoption - Core practices

	(1) Core index	(2) Suitable seed	(3) Seed replacement	(4) Seed treatment	(5) Pesticide	(6) Herbicide	(7) Fertilizer dap	(8) Fertilizer mop	(9) Fertilizer urea	(10) Line method	(11) Micronutrient zinc
Panel A: Year 1 - Kharif 21											
Treated	0.068**	-0.001	0.017*	0.010	0.025*	0.009	0.004	0.003	0.018*		
	(0.028)	(0.013)	(0.009)	(0.014)	(0.014)	(0.009)	(0.015)	(0.013)	(0.009)		
N	4057	4057	4057	4057	4057	4057	4057	4057	4057		
R^2	0.016	0.005	0.007	0.012	0.009	0.010	0.012	0.010	0.011		
Control mean	0.003	0.201	0.883	0.290	0.420	0.077	0.490	0.730	0.877		
Baseline mean		0.183	0.800	0.088			0.396	0.652	0.863		
Panel B: Year 2 - Kharif 22											
Treated	0.048	-0.013	0.015	0.020	0.020	0.015	-0.004	0.014	0.019	-0.004	-0.001
	(0.039)	(0.017)	(0.012)	(0.017)	(0.017)	(0.010)	(0.015)	(0.018)	(0.018)	(0.014)	(0.008)
N	2910	2910	2910	2910	2910	2910	2910	2910	2910	2908	2910
R^2	0.019	0.010	0.011	0.010	0.009	0.012	0.014	0.025	0.041	0.008	0.015
Control mean	-0.000	0.344	0.887	0.253	0.389	0.073	0.494	0.510	0.519	0.170	0.049
Baseline mean		0.185	0.796	0.091			0.393	0.649	0.851	0.146	0.023

Note: Data are from the midline and endline survey of Cohort 1 farmers. The dependent variables are dummy variables indicating whether farmers adopted the specific practice correctly or not. The selected practices are practices that are relevant to all farmers and constitute the "core" summary index. The analysis sample is the full set of farmers who consented to the respective survey and responded to all relevant adoption questions. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). Fixed effects at the randomization strata level are included. Robust standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table A5: The impact of treatment on adoption - Transplanting practices

	(1) Transplanting index	(2) Nursery fertilizer urea	(3) Suitable seed*	(4) Seed replacement*	(5) Seed treatment*	(6) Pesticide*	(7) Herbicide*	(8) Nursery fertilizer fym	(9) Nursery fertilizer dap	(10) Nursery fertilizer mop	(11) Fertilizer dap*	(12) Fertilizer mop*	(13) Fertilizer urea*	(14) Micronutrient zinc*	(15) Line method*	(16) Transplanting time
Panel A: Year 1 - Kharif 21																
Treated	0.096***	0.002	0.008	0.021**	0.011	0.027*	0.008	0.010	0.021	0.023	-0.000	0.002	0.013	0.018*		
	(0.031)	(0.014)	(0.013)	(0.009)	(0.016)	(0.015)	(0.009)	(0.011)	(0.016)	(0.016)	(0.016)	(0.013)	(0.009)	(0.010)		
N	3458	3458	3458	3458	3458	3458	3458	3458	3458	3458	3458	3458	3458	3458		
R^2	0.015	0.009	0.007	0.006	0.014	0.008	0.009	0.006	0.007	0.004	0.013	0.011	0.009	0.011		
Control mean	0.005	0.742	0.202	0.893	0.313	0.419	0.076	0.597	0.334	0.393	0.473	0.738	0.887	0.081		
Baseline mean			0.182	0.828	0.092						0.385	0.657	0.863	0.027		
Panel B: Year 2 - Kharif 22																
Treated	0.109**		-0.001	0.022*	0.017	0.016	0.008	0.041**	0.001	0.048***	0.007	0.012	0.029	-0.002	-0.006	0.026
	(0.042)		(0.019)	(0.012)	(0.020)	(0.019)	(0.011)	(0.018)	(0.017)	(0.017)	(0.015)	(0.019)	(0.020)	(0.009)	(0.015)	(0.020)
N	2454		2454	2454	2454	2454	2454	2454	2454	2454	2454	2454	2454	2454	2454	2454
R^2	0.018		0.012	0.009	0.011	0.007	0.010	0.016	0.009	0.016	0.013	0.022	0.046	0.014	0.008	0.008
Control mean	0.001		0.340	0.893	0.280	0.390	0.077	0.683	0.272	0.245	0.500	0.527	0.537	0.054	0.178	0.541
Baseline mean			0.183	0.829	0.095						0.376	0.661	0.854	0.026	0.147	0.766

Note: Data are from the midline and endline survey of Cohort 1 farmers. The dependent variables are dummy variables indicating whether farmers adopted the specific practice correctly or not. The selected practices are practices that are relevant to farmers that practice transplanting and constitute the "transplanting" summary index. Practices with "*" in the column titles are included in creation of the "core" summary index. The analysis sample is the set of farmers who consented to the respective survey, responded to all relevant adoption questions, and reported transplanting in the Kharif 2020-2021 season in the baseline survey. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), whether the survey was completed after switching the modality, and whether the endline survey was conducted in the second batch (for endline outcomes only). Fixed effects at the randomization strata level are included. Robust standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table A6: The impact of treatment on rice production cost and investment (Year 1)

	Variable costs (Rs)									Investments (Rs)		
	(1) Seed	(2) Fertilizer	(3) Pesticide	(4) Irrigation	(5) Worker	(6) Tractor	(7) Transport	(8) Other	(9) Total	(10) Irrigation	(11) Tractor	(12) Total
Panel A: Indicators for incurred specific expenditure												
Treated	0.009	0.010*	0.025	0.017	0.023*	0.028***	0.021	0.030*		0.009	0.013*	0.017
	(0.020)	(0.005)	(0.022)	(0.019)	(0.013)	(0.010)	(0.021)	(0.018)		(0.010)	(0.007)	(0.011)
N	1929	1929	1929	1929	1929	1929	1929	1929		1929	1929	1929
R^2	0.106	0.022	0.099	0.042	0.073	0.012	0.016	0.012		0.008	0.015	0.017
Control mean	0.696	0.973	0.459	0.251	0.895	0.941	0.737	0.196		0.046	0.023	0.068
Panel B: Values of specific expenditure												
Treated	68.227	237.727	33.336	3.561	371.015	142.038	78.724	21.460	796.483	133.862**	199.078*	381.531***
	(51.472)	(155.585)	(26.129)	(24.408)	(270.247)	(198.266)	(61.355)	(18.343)	(571.580)	(67.034)	(108.114)	(140.433)
N	1929	1929	1929	1929	1929	1929	1929	1929	1929	1929	1929	1929
R^2	0.208	0.290	0.154	0.038	0.360	0.070	0.060	0.030	0.455	0.012	0.014	0.020
Control mean	1249.552	5175.614	376.541	261.267	9139.174	6926.658	1321.238	162.189	25634.269	138.181	264.700	440.971
Control sd	1407.140	4162.338	578.825	570.972	8459.240	5322.160	1449.391	400.715	18845.093	1036.666	2304.576	2710.573

Note: Data are from the midline survey of Cohort 1 farmers. The dependent variables are (Panel A) dummy variables indicating whether farmers spent on specific inputs, and (Panel B) continuous variables indicating the amount of money (INR) that farmers spent on specific inputs, which are winsorized at the 95th percentile for variable costs and at the 99th percentile for investments. Everyone except 9 farmers (99.6%) has incurred some variable costs, so we do not examine the treatment impact on this dummy variable (Panel A Column (9)). All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), and whether the survey was completed after switching the modality. Fixed effects at the randomization strata level are included. Robust standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table A7: The occurrence of weather shocks

Weather event	Month	% of rural households in Odisha affected	% of study sample farmers affected
Panel A: Year 1 - Kharif 2021			
Excess rainfall	August-October	21%	30%
Cyclone	December	29%	46%
Panel B: Year 2 - Kharif 2022			
River flooding	August	/	17%
Scarce rainfall	August-October	43%	33%

Note: Weather shocks are defined as follows: "excess rainfall" refers to received more than 250 mm rainfall over two days in the period of August 16 to October 15 in 2021 (i.e., the main growing stage) and is defined at the block level; "cyclone" refers to received more than 50 mm rainfall over two days in the period of December 4 to December 10 in 2021 (i.e. the period of Cyclonic Storm Jawad) and is defined at the block level; "river flooding" refers to experienced river flooding in August 2022 and is defined at the panchayat level; and "scarce rainfall" refers to received less than 450 mm rainfall over the period of August 16 to October 15 in 2022 (i.e., the main growing stage) and is defined at the block level.

Table A8: The impact of treatment on rice revenue and profit (Year 1) [alternative outcome construction]

	Full sample	In-person sample				
	(1)	(2)	(3)	(4)	(5)	(6)
	Harvest (kg)	Harvest (kg)	Reported revenue (Rs)	Value of harvest (Rs)	Variable cost (Rs)	Imputed profit (Rs)
Panel A: Profit alternative version 1						
Treated	76.487*	107.991	656.696	1455.572	893.612*	475.784
	(0.094)	(0.108)	(0.460)	(0.141)	(0.064)	(0.576)
		[0.381]	[0.820]	[0.381]	[0.381]	[0.820]
N	3835	1929	1929	1929	1929	1929
R^2	0.395	0.422	0.355	0.418	0.486	0.155
Control mean	2374.902	2238.218	16606.245	32774.124	24612.232	8057.171
% Change	3.22	4.82	3.95	4.44	3.63	5.91
Panel B: Profit alternative version 2						
Treated	76.487*	107.991	656.696	1927.329	893.612*	975.047
	(0.094)	(0.108)	(0.460)	(0.195)	(0.064)	(0.462)
		[0.379]	[0.752]	[0.431]	[0.379]	[0.752]
N	3835	1929	1929	1929	1929	1929
R^2	0.395	0.422	0.355	0.369	0.486	0.201
Control mean	2374.902	2238.218	16606.245	50544.226	24612.232	25855.563
% Change	3.22	4.82	3.95	3.81	3.63	3.77

Note: Data are from the midline survey of Cohort 1 farmers. The dependent variables are (1) and (2) total rice harvest (kg), (3) self-reported revenue from rice sales (INR), (4) total value of rice harvest (INR), (5) total variable cost of rice cultivation (INR), and (6) imputed profit of rice harvest (total value of harvest minus total variable cost, INR). Values of harvest are constructed using sales prices in Panel A and using both retail prices and sales prices in Panel B. Variable costs are summation of winsorized cost items in Panel A and B. Profits are constructed using the value of harvest and variable cost in the same panel. Harvest, revenue, value of harvest, and cost outcomes are winsorized at the 95th percentile, and imputed profit outcome is winsorized at the 2.5th and the 97.5th percentiles. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), and whether the survey was completed after switching the modality. Fixed effects at the randomization strata level are included. Robust standard errors are calculated: * p<0.10, ** p<0.05, *** p<0.01. Their corresponding p-values are reported in parentheses, and Romano and Wolf's step-down adjusted p-values are reported in brackets. % changes represent the treatment impacts over the control means.

Table A9: The treatment impact in areas with and without a weather shock (Year 1) [alternative outcome construction]

	Imputed profit (Rs)	
	(1) Alternative version 1	(2) Alternative version 2
Panel A: Excess rainfall		
Treated	-244.373	60.807
	(0.703)	(0.952)
	[0.998]	[1.000]
Treatment × Excess rainfall	2748.700**	3489.480
	(0.023)	(0.119)
	[0.204]	[0.723]
N	1929	1929
Ctrl. mean (No excess rainfall)	10415.732	28187.468
Ctrl. mean (Excess rainfall)	1101.414	18978.419
Total effect for excess rainfall areas	2504.327**	3550.287*
	(0.013)	(0.080)
	[0.202]	[0.581]
Panel B: Cyclone		
Treated	111.163	871.989
	(0.920)	(0.633)
	[1.000]	[0.998]
Treatment × Cyclone	729.408	206.164
	(0.557)	(0.916)
	[0.996]	[1.000]
N	1929	1929
Ctrl. mean (No cyclone)	12081.677	30374.457
Ctrl. mean (Cyclone)	3997.971	21297.713
Total effect for cyclone areas	840.571	1078.152
	(0.208)	(0.250)
	[0.914]	[0.948]

Note: Data are from the midline survey of Cohort 1 farmers. The dependent variables are imputed profits of rice harvest, and winsorized at the 2.5th and the 97.5th percentiles. Weather shocks are defined as follows: "excess rainfall" refers to received more than 250 mm rainfall over two days in the period of August 16 to October 15 in 2021 (i.e., the main growing stage) and is defined at the block level; "cyclone" refers to received more than 50 mm rainfall over two days in the period of December 4 to December 10 in 2021 (i.e. the period of Cyclonic Storm Jawad) and is defined at the block level. All regressions control for vegetation indices from the two pre-intervention Kharif seasons (2019 and 2020) and demographic characteristics collected in the baseline — the respondent's gender, age, literacy level, sole ownership of the primary phone, and household wealth level, access to irrigation in the upcoming Kharif 2021 season, and whether the respondent's primary phone is a feature phone — and imputation dummies for missing baseline measures. Additionally, we control for whether the respondent completed the baseline compensation survey before the start of the intervention, the survey modality assignment for follow-up surveys (phone or in-person), and whether the survey was completed after switching the modality. Two additional variables - total rice farming area reported at the baseline and whether its value is imputed with block median - are controlled. Fixed effects at the randomization strata level are included. Standard errors are clustered at block level: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Their corresponding p-values are reported in parentheses and Romano and Wolf's step-down adjusted p-values are reported in brackets.

B Assignment of Survey Modality

We used in-person and phone surveys in the two follow-up surveys with cohort 1 farmers.

For the midline follow-up survey, the survey modality assignment followed a two-stage randomization. We first randomly selected 40% of panchayats as the "phone survey only" areas and the rest of panchayats as the "mixed modes of survey" areas, stratifying by district. We then randomly assigned 50% of farmers in the mixed modes of survey areas to receive an in-person survey, and the remaining farmers to receive a phone survey. To maximize the response rate, we switched the mode of survey after three unsuccessful attempts.

For the endline survey, we randomized farmers into four groups, using a 2×2 factorial design and exogenously varied survey modality (phone or in-person) and the intensity with which we seek to contact individuals (regular or intensive). Specifically, we randomly allocated farmers in the treatment and control groups to one of four survey protocols (allocation fractions in parentheses):

- S1 (10%): We contacted each respondent up to 4 times in-person.
- S2 (40%): We contacted each respondent up to 2 times in-person. After two failed in-person attempts, we followed up with up to 4 attempts to reach respondents by phone.
- S3 (10%): We contacted each respondent up to 6 times by phone.
- S4 (40%): We contacted each respondent up to 4 times by phone. After 4 failed phone attempts, we followed up with up to 2 attempts to reach each respondent in person.