Misspecified Learning in Technology Adoption: Experimental Evidence from Fertilizer Use in China

Wei Lin* Binkai Chen† Ao Wang‡

November 12, 2021§

Click here for the most recent version

Abstract

This paper investigates agents' simultaneous learning about multiple interacting technologies in the context of fertilizer application in China. We first present experimental evidence that farmers overuse nitrogen fertilizers and underuse phosphorus and potassium fertilizers, relative to the personalized fertilizer recommendations based on plot-level soil analysis. Our first-phase intervention that provides customized fertilizer recommendations leads to reduced nitrogen applications and increased phosphorus/potassium uses. Average yields and revenues are 5-7% higher, while total fertilizer costs remain unchanged. These results are also consistent with a meaningful reduction in greenhouse gas (N_2O) emissions linked to nitrogen overuse. Survey data suggest that farmers overestimate the return to nitrogen because it produces a salient signal on crops by increasing greenness, but they underestimate the effectiveness of phosphorus and potassium because their effects are barely observable during the growing stages. Motivated by these facts, we then propose a model of misspecified learning in which agents face two technologies with unknown returns. In learning about the effectiveness of both technologies, the overestimation of the return to the first technology causes an undervaluation and underuse of the second technology. To further test the model, we design a second-phase intervention that distributes leaf color charts to farmers to correct their overestimation of the return to greenness. Consistent with the model prediction, the intervention not only reduces farmers' nitrogen use immediately, but also induces gradual learning of phosphorus and potassium; the proportion of farmers using phosphorus and potassium both increase by 6 percentage points, relative to 4% and 9% in the control group.

Keywords **: Mislearning, Overestimation, Effectiveness, Undervaluation**

JEL classification: **D83, O13, O33, Q16, Q51**

*Corresponding author. UC Berkeley. Email: wei_lin@berkeley.edu.

†Central University of Finance and Economics. Email: chenbinkai@cufe.edu.cn.

‡UC Berkeley. Email: ao.wang@berkeley.edu.

§We are incredibly thankful for continued support and advice throughout this entire research from Marco Gonzalez-Navarro, Benjamin Faber, Jeremy Magruder, and Adam Szeidl. For very supportive and helpful comments, we would like to thank Sam Asher, Alain de Janvry, Supreet Kaur, Ethan A. Ligon, Aprajit Mahajan, Edward Miguel, Elisabeth Sadoulet, Brian Wright, and Yuen Ho, as well as numerous seminar participants at UC Berkeley. We gratefully acknowledge generous funding and support by the Weiss Family Fund and CEGA. Binkai Chen acknowledges support from the Social Science Foundation of China. All remaining errors are our own. We received IRB approval from UC Berkeley. AEA RCT Registry- 0006477.

1 Introduction

Standard learning models assume agents converge to the optimal usage of technology through learning and observation, which fail to explain ample field evidence that technologies can be incorrectly used even after many years of learning.[1] Furthermore, the learning problem becomes considerably more complicated when agents attempt to learn about multiple technologies simultaneously, as it requires agents to understand the returns to all of the technologies. A seminal paper by Hanna, Mullainathan, and Schwartzstein (2014) explores the role of mislearning in modeling the frictions caused by incorrect use of technologies. Yet we still know little about decision-making in the adoption of multiple technologies and the underlying mechanisms that can induce potential overuse and underuse.

This paper sheds light on this multi-technology learning problem by explicitly modeling the interactive effects that occur when agents learn about two technologies simultaneously and presents novel experimental evidence from the utilization of fertilizer technology in China. We chose this context for two reasons. First, fertilizers, which consist of multiple technologies (nitrogen, phosphorus and potassium), are essential to productivity growth in most developing countries but are often underused or overused.[2] Both the underuse and overuse of fertilizers are especially puzzling because farmers tend to engage with these technologies over decades, which should be a sufficient time frame for them to learn the correct usage according to classic learning-by-doing models. Second, the efficiency loss and greenhouse gas emissions (CO_2 and N_2O) caused by the misuse of fertilizers, particularly the overuse of nitrogen, are especially pronounced in China. According to the International Atomic Energy Agency (IAEA), agricultural activities, mainly the use of fertilizers, contribute approximately 10 -30% of global greenhouse gas emissions,[3] driven mostly by activities in Brazil, Europe, India, the US, and particularly China, which accounts for 28% of global fertilizer use.

Using administrative data from soil analysis of millions of local plots conducted by the government, we first observed that farmers simultaneously overuse nitrogen (N) fertilizer in the growing stage and under use phosphorus (P) and potassium (K) fertilizers throughout the cropping cycle. To study whether farmers can correctly learn the effectiveness and optimal application levels of different fertilizers, we designed and implemented a two-phase randomized controlled trial (RCT) among 1,200 farmers in 200 villages. In the first phase of our experiment, we provided farmers with customized fertilizer recommendations based on the soil analysis at the plot level. We randomly varied whether farmers only received the soil testing ($T1$);[4] whether they received the soil analysis and customized and dynamic fertilizer recommendations through a smart mobile application

[1] For instance, agents may overuse technologies such as antibiotics and pesticides but under-invest in firm management, health products, and agricultural inputs.

[2] For example, Kenyan farmers underuse fertilizers due to procrastination (Duflo, Kremer, and Robinson, 2008, 2011), Chinese farming households overuse fertilizers as a consequence of low education (Cui et al., 2018), and Mexican maize farmers misuse different types of fertilizers because of soil heterogeneity (Corral et al., 2020).

[3] This share varies among different countries; for example, agriculture accounts for 10% of total U.S. greenhouse gas emissions (EPA).

[4] Specifically, farmers received detailed information about the soil quality and micronutrient content of their plots.

($T2$); whether they received the soil analysis, the mobile application, and a training session from agricultural extension specialists ($T3$); or whether they served as a pure comparison group. In the $T3$ group, the extension meetings were conducted one-on-one in-person, where specialists showed the experimental effects of phosphorus (P) and potassium (K) on yields to update farmers' beliefs about the effectiveness of these fertilizers. The randomization was carried out at the village level.

The results from the first-phase experiment indicate that the treatments that combine soil analysis data with customized support (T_2 and T_3) significantly reduced farmers' application of nitrogen fertilizer and increased use of phosphorus and potassium fertilizer relative to the control group. Specifically, we find that $T2$ (App) and $T3$ (App + in-person training) effectively reduced nitrogen (N) fertilizer use by 3.92-4.43 kg per mu [5] at the intensive margin, roughly 13.3-14.7% of the control mean ($p < 0.01$). This substantial reduction in nitrogen fertilizer use occurred mainly during the growing stage of the crops.

For the other fertilizers, the overall usage of phosphorus in the $T2/T3$ group increased significantly by 2.34/2.72 kg per mu ($p < 0.01$). Specifically, the extensive margin substantially contributed to this increase: the proportion of households using top-dressing phosphorus fertilizer jumped by 23.2-23.9 percentage points ($p < 0.01$), compared to a control mean of 3 percent. Similarly, for potassium we find usage increased by 1.37/2.89 kg per mu in the $T2/T3$ group, again mainly driven by the extensive margin, as there was a 24.1-32.4 percentage point increase in the proportion of households using top-dressing potassium fertilizers. These large extensive-margin impacts suggest that the treatments initiated farmers' experimentation with phosphorus (P) and potassium (K) fertilizers, since they seldom applied P and K fertilizers prior to the experiment. In contrast, there were no significant effects on fertilizer usage in the $T1$ group, likely because farmers did not understand how to use the raw soil testing data to inform their farming practices.

How large was the inefficiency from the misuse of different fertilizers? We then explore the changes in yields/profits due to the re-optimization of different fertilizers. The change in fertilizer use caused by $T2$ and $T3$ led to a significant increase in yields by 5-7% and revenues by 6.0-6.9%, without changing the cost of fertilizers and other inputs. These results are also consistent with a meaningful reduction in greenhouse gas (mainly N_2O) emissions linked to the excessive use of nitrogen.

Survey evidence suggests that farmers overestimate the return to nitrogen fertilizer because it produces a salient signal on crops by increasing greenness, but they underestimate the effectiveness of phosphorus and potassium fertilizers, which increase yields but have few immediately observable impacts on crops. This simultaneous and persistent overuse and underuse of different fertilizers cannot be explained by standard Bayesian learning or the selective attention in Hanna, Mullainathan, and Schwartzstein (2014). To study the mechanisms, we propose a model of misspecified learning, building on work by Heidhues, Kőszegi, and Strack (2021), in which agents simultaneously learn about two different types of interacting technologies with unknown returns.

In our model, agents face two technologies and overestimate the return to the first technology.

[5] 1 mu = $\frac{1}{15}$ hectare. 3.92-4.43 kg per mu is approximately equivalent to 60 kg per hectare.

The gap between their subjective beliefs and observed profits is rationalized by lower-than-true perceptions of the effectiveness of phosphorus and potassium. As a result, the overestimation of the return to greenness distorts agents' actions in the use of the first technology; such distorted actions influence their valuation of the effectiveness of the second technology. Consequently agents overuse the first and under use the second technology. Our model generates three main predictions: 1) farmers are trapped in a sub-optimal equilibrium, where nitrogen is overused and phosphorus and potassium are undervalued and underused; 2) farmers' perception about the effectiveness of phosphorus and potassium moves toward the true value as their actions in fertilizer use move to the optimal level; and 3) correcting the overestimation of the return to greenness not only leads to immediate learning about nitrogen, but also induces farmers to learn more about phosphorus and potassium fertilizers. Among them, prediction (1) is verified by the findings from the first-phase experiment.

Next, we present evidence consistent with model prediction (2) regarding beliefs about the effectiveness of different fertilizers. Our experimental results from the first phase shows that farmers' valuation of the effectiveness of nitrogen remained unchanged since most farmers (95.2% in the control group) already understood the effects of nitrogen (N) on greenness. We then explore $T2$ intervention which encouraged farmers to change their actions in fertilizers, but didn't intervene directly to change their beliefs. We find 22.0, 20.8, and 17.1 percentage points more farmers correctly understood the relationship between P/P/K and flower timing/root length/ grains' density after the $T2$ intervention. In the T_3 group, these effects on beliefs about the effectiveness of different fertilizers were almost doubled due to the presence of farmers' social learning from agricultural extension specialists.[6]

Can model-driven interventions help resolve learning failure in the application of different fertilizers? To test model prediction (3), in the second-phase field experiment, we randomly varied whether farmers received leaf color charts (LCC) to help them better calibrate the optimal level of nitrogen to be used on their crops. The goal of the LCC intervention was to correct farmers' overestimation of the return to greenness in the production function. By following the user instructions, farmers could compare the actual leaf color of their crops to the greenness on the charts to make informed decisions about optimal top-dressing fertilizer applications.

The results from the second-phase experiment confirm our model prediction (3) that correcting misperceptions about one technology (nitrogen fertilizer) induces agents to learn about and experiment more with the second technology (phosphorus and potassium fertilizers). Specifically, the leaf-color-chart intervention immediately reduced farmers' nitrogen fertilizer application by 3.76 kg per mu, corresponding to a 12.3% decrease compared to the control mean. The leaf color charts also encouraged a small proportion of farmers to experiment with using phosphorus (6.62 percentage points) and potassium (6.66 percentage points) for the first time. The results suggest that reducing farmers' misspecification in one dimension allowed them to learn about the other dimensions.

[6]By contrast, in T2, we just provided the mobile application and instructions to farmers; thus, their updating of beliefs should have only come from self-learning.

Using two-stage least squares, we estimate that the changes in fertilizer use caused by the leaf color charts led to a 3.4% increase in average yields and a 4% increase in revenues compared to the control group.

We then discuss some issues with measurement problems and alternative theories for nitrogen overuse. One natural concern is that the self-reported inputs and outputs may affect the results. To address this concern, we collected fertilizer usage in three different ways, both from the aggregate perspective (amount of use in a year) and in multiple stages (amount of use in different growing stages). We find that data in these questions are quite consistent, which makes misreporting issues unlikely. Another concern is that supply-side sellers and price may affect farmers' decision-making. We present direct and indirect evidence that these factors didn't really drive the overuse of nitrogen fertilizers. First, the farmers' bias with regard to fertilizer application concentrates on the growing stages. Second, their decision-making is mostly based on the greenness signals during the growing season and they believe that the greener the batter.

Taken together, the new mechanisms proposed by this paper are externally relevant to other contexts in which agents learn about multiple technologies simultaneously.[7] Our results also demonstrate that cost-effective interventions guided by theory can correct agents' sub-optimal input choices, with important policy implications. A cost-benefit analysis indicates that the profit gains exceed the costs for the app-based interventions (T_2), the extension services intervention (T_3), and the leaf color chart intervention (LCC). However, there is a trade-off of cost versus speed of realized results for policymakers when choosing between these interventions. On the one hand, the app-based intervention (T_2) and extension services intervention (T_3) allow agents to re-optimize input choices immediately, but they are more costly to implement and require plot-level soil testing data. In comparison, the leaf color chart intervention is more easily scaled due to significantly lower costs, but induces slower learning. We also analyze the aggregate benefits of these interventions. A back-of-the-envelope estimate suggests that total greenhouse gas emissions would reduce by 37.4 million tons per year (0.4% of China's annual CO_2 emissions), while rice farmers' revenues would increase by roughly 30 billion RMB.

Our work contributes to three main strands of literature. Our research questions are most related to topics on technological learning and misuse of a single technology. The existing theories and empirical evidence in the field explore the under-investment in agricultural technologies due to the cost of labor (Foster and Rosenzweig, 2010), procrastination (Duflo, Kremer, and Robinson, 2011), distance to public transport (Suri, 2011), and low attention to one particular input dimension (Hanna, Mullainathan, and Schwartzstein, 2014).[8] We complement this strand of literature

[7]For example, farmers face trade-off between pesticides and bug repellent variety of seed for crops. Agents face trade-off between taking antibiotics and improving hygiene conditions. In these two examples, the first technology generates relatively more observable feedback, while the second technology has less salient impacts.

[8]Beyond the agriculture sector, recent studies have looked at the misuse of technology in other fields, such as deworming (Miguel and Kremer, 2004; Hamory et al., 2021), antibiotics (Currie, Lin, and Meng, 2014), vaccines (Karing, 2018), new health products (Dupas, 2014); management (Bloom et al., 2013) and new products in firms (Atkin et al., 2017); energy efficiency products Allcott and Taubinsky (2015) and energy-saving stoves (Berkouwer and Dean, 2019).

in two aspects: 1) We first experimentally document the existence of simultaneous overuse and underuse of different technologies when more than one technology is at play;[9] 2) We propose a new mechanism—mislearning between different technologies. Our model and survey evidence suggest that the misperception of one technology can influence the perception of the effectiveness of other technologies.

Our methodology and results on mechanisms build on recent theoretical literature on misspecified learning (Heidhues, Kőszegi, and Strack, 2018; Fudenberg, Lanzani, and Strack, 2021; Heidhues, Kőszegi, and Strack, 2021). The key intuition of these theories is that misspecification in the production function affects agents' actions, and such distorted actions then change agents' valuation of the technology. We contribute to this work by extending the model to two dimensions/technologies, and studying mislearning transmission between two technologies. Our paper, to the best of our knowledge, is the first to experimentally test the theory of misspecified learning in the field. We provide evidence consistent with the model predictions and find that a theory-based intervention in our second-phase experiment indeed can resolve learning failure in fertilizer application.

Our interventions and policy implications leverage information communication technology (ICT) to improve farmers' efficiency and reduce environmental damage and greenhouse gas emissions. The role of ICT has been considered by several studies in agriculture (Casaburi et al., 2014; Casaburi, Kremer, and Ramrattan, 2019; Fabregas, Kremer, and Schilbach, 2019; Cole and Fernando, 2021), in environmental protection (Greenstone et al., 2020), and in firm and business performance (Jensen, 2007; Jensen and Miller, 2017). Our paper relates to this literature by providing farmers with customized agricultural services through a smart mobile application. This is particularly relevant in low- and middle-income countries (LMICs) since many smallholder farmers in LMICs lack access to science-based agricultural advice. In these countries, information provision to farmers is often "top-down" and not localized, which results in inadequate diagnosis of farmers' needs with respect to local agro-ecological settings and diverse farm-level characteristics. By taking advantage of administrative data on soil testing, our mobile application serves as a precise benchmark for optimal fertilizer application and thus reliably complements the service provided by extension agents. This paper also adds to the studies on greenhouse gas emissions (Gilbert, 2012; Tian et al., 2020). Fertilizer pollution in agriculture is often under-evaluated in economics literature but has an unignorable impact on the global environment.

The paper proceeds as follows. Section 2 describes the setting and first-phase experimental design. Section 3 discusses results from the first-phase field experiment. Section 4 presents our theoretical framework. Section 5 details the second-phase experimental design and interprets the results. Section 6 concludes.

[9] We compare our impacts to related interventions. In terms of nitrogen-fertilizer application, Chen et al. (2014) estimate that an integrated soil–crop system management (ISSM)-based recommendation led to higher yields (18–35%) and a reduction in nitrogen usage (4–14%).

2 Setting and First-phase Experimental Design

2.1 Fertilizer Consumption: China vs. the World

As a result of modern technological advances, fertilizer has been able to generate high returns for farmers, and it was responsible for significant growth in agricultural yields during the 20th century. With a considerable increase in the use of fertilizer, traditional agricultural extension has not always provided the most useful instructions for farmers or given them knowledge about alternative farming practices (Beaman et al., 2021), especially in less developed countries, where small-landholding farmers may be ill-informed regarding safe and sustainable fertilizer use standards due to a lack of extension support. This is the case in China, where 98% of farming households have a farming plot of less than 2 hectares (Wu et al., 2018).

Excessive use of fertilizers, especially nitrogen, results in massive emissions of greenhouse gases (CO_2, N_2O). As shown in Figure A1 in Appendix A, developed countries such as the US, Australia, Germany, the Netherlands, and Denmark, and developing countries including India, Bangladesh[10], and China are all experiencing excessive nitrogen application. Therefore, efficient use of fertilizer is essential for sustainable development and the fight against global warming (Tian et al., 2020).

There is a longstanding discussion on the overuse of nitrogen fertilizer in China.[11] China is the world's largest fertilizer producer and consumer, accounting for roughly 28.8% of global fertilizer use, while its arable land is only 7.6-9% of the world's total. Figure 1 shows the fertilizer application intensity (kg/ha) during 1961-2014 among several countries, including Brazil, Bangladesh, China, Germany, India, Kenya, Mexico, and the United States. China has had a rapid increase in fertilizer use, especially after 1980 due to the introduction of synthetic fertilizers. In developed countries such as Germany, usage reached a peak around 1980, and dramatically decreased after 2000,[12] when the German government started to impose restrictions on fertilizer use. For developing countries such as Brazil and Bangladesh, fertilizer applications follow a pattern similar to China at an earlier stage of its development.

2.2 Design of Interventions

A recent study on fertilizer application in China suggests that a reduction of 30% to 50% in the application of fertilizer would not necessarily compromise yields (Cui et al., 2018). To figure out whether farmers in China apply a sub-optimal mix of fertilizers and understand the economic consequences of such misuse, we designed and implemented an experiment consisting of two main phases: (i) change farmers' actions in choosing the level of fertilizer application by providing precise soil analysis and customized fertilizer recommendations; (ii) correct farmers' overestimation of the return to greenness by distributing leaf color charts (LCC) among 1,200 farmers in 200 villages in Leiyang, Hunan province.

[10] See Rahman and Zhang (2018); Islam and Beg (2021).

[11] See discussions on the relationship between excessive nitrogen use and yield by Chen et al. (2014); Zhang et al. (2015); Cui et al. (2018); Zhang et al. (2015); Wu et al. (2018), etc.

[12] We ignore the period around 1990 since the statistics might also experience changes due to reunification.

Universal Soil Testing Program. To provide farmers with customized recommendations, we first needed administrative data on soil analysis from the universal soil testing program implemented by Hunan province. Figure 2a shows the distribution of universal soil testing programs in our experimental site. Each green dot is a testing point/paddy field where an agricultural extension specialist collected a soil sample and analyzed its micronutrient components in the laboratory. Similar to the soil analysis protocol in Corral et al. (2020), the project recorded (for each green dot) the soil texture (type of soil under soil taxonomy), pH levels, levels of the primary macronutrients (nitrogen (N), phosphorus (P), and potassium (K)) and secondary macronutrients (calcium, magnesium, and sulfur), and the level of organic matter.[13]

Our baseline survey in April 2020 shows that the dissemination of the testing results to farmers was quite low. Less than 5% of farmers among 1,200 surveyed households had acquired this information. Two obstacles prevented dissemination: 1) the testing results are too abstract and unreadable from the farmers' perspective; and 2) the cost of face-to-face dissemination is high.

Recommendations. To address obstacle (1), in addition to the soil analysis, farmers need the corresponding fertilizer recommendations.[14] The customized recommendations of fertilizer dosage are generated based on three elements: a reliable production function of different fertilizers, crop-specific micronutrient demand, and price to maximize farmers' profits. The primary production function was simulated based on the results from 10,000 experimental trials in Hunan province, in which three fertilizers (nitrogen, phosphorus and potassium) were randomly applied at four different levels, 1) *zero*, 2) *0.5 times local average recommendations*, 3) *local average recommendations*, and 4) *1.5 times local average recommendations*.[15] The micronutrients needed are calculated as a function of micronutrients in the soil and targeted outputs multiplied by the average price.

Tools for Intervention: A Smart Mobile Application. To address obstacle (2) and disseminate individual testing results and customized fertilizer recommendations more effectively, we partnered with local governments in Hunan province and a technology company,[16] co-developing a mobile application that has the following appealing features: 1) It is fully endorsed by the Hunan and Leiyang governments (Figure A2a) and can provide support for up to 15 crops (Figure A2b). It connects the farmer to the administrative dataset on the universal soil testing results for millions of plots in Hunan province. Through GPS tracking (Figure A3a) or by selecting the region (Figure A3b) , farmers can automatically acquire the soil analysis results of the nearest testing plot where a soil sample was collected and analyzed. 2) The smart mobile application displays the

[13]See similar protocols of soil testing in Fishman et al. (2016); Murphy et al. (2020); Harou et al. (2018), etc.

[14]The recommendations are not weather- and temperature-specific. The algorithm is under normal rain and temperature conditions.

[15]China officially called such experimentation the 3-4-14 trials. The number 3 indicates three different fertilizers, the number 4 means four different levels of the application of different fertilizers, and the number 14 indicates 14 different combinations and trials. See the official announcement by the Ministry of Agriculture and Rural Affairs of China: http://www.moa.gov.cn/govpublic/CWS/201405/t20140523_3915330.htm and http://m.ynforestry-tec.com/upload/manager/image/201908/21/20190821092521548230938.pdf

[16]Tianjiandao technology software company. This mobile application was first developed in 2015. But our baseline survey shows that the dissemination of soil analysis data to farmers is very low in our experimental site. We joined in 2019 to improve the algorithms, profit generation, and fertilizer recommendations. The latest updated version was on 08/27/2020.

recommended and dynamic combination of different individual fertilizers (N-fertilizer, P-fertilizer, and K-fertilizer) to be used in each stage (planting and growing stages), as shown in Figure A4a. 3) Since most farmers are using N-P-K compound fertilizer, it also recommends the optimal mix of N-P-K compound fertilizer and individual fertilizers (N/P/K) in a dynamic timeline. As shown in Figure A4b in Appendix A, it displays the amount of N-P-K compound fertilizer and nitrate fertilizer (N) to be used before the planting stage (the basal fertilizer stage) and the amount of top-dressing nitrogen, phosphorus, and potassium fertilizers to be used during the growing stage (the top-dressing stage).[17]

2.3 Experimental Timeline and Framework

Randomization. We first acquired a full list of 348 villages from local Department of Agriculture, with basic information on total land area, average pH value, average N/P/K, and organic matter in the village-level plots. In order to screen villages with sufficient rice farmers, we dropped villages which had a total agricultural land area less than 1000 mu (equivalent to 66.7 hectares). We then listed these villages by their alphabetical order and randomly selected 200 villages using the random number generator. Next, as shown in Figure 3, we randomized these 200 villages into four arms (T_1, T_2, T_3 and Control).[18] In each village, we randomly select six rice farming households based on the list provided by village head. To summarize:

1) T_1: ST group. In this group, 300 farming households in 50 villages were provided with individualized soil testing analysis data only. Farmers were informed of the level of micronutrients, including nitrogen/ phosphorus/ potassium, in their plots.

2) T_2: App group. In this arm, 300 farming households in 50 villages were provided with the mobile application and detailed instructions by a well-designed handbook and instructive video, as shown in Appendix C. To ensure farmers or their household members understood the mobile application well, we also asked them to repeat the procedures for use during the visit. Enumerators recorded the whole process in the survey.

3) T_3: App + In-person agricultural extension agents' training (AEA's training). In this group, 300 farmers were not only offered the smart mobile application, but also given the agricultural extension services. During the visits, the agricultural extension agents held an in-person and one-to one training session, showing the experimental relationship between phosphorus (P)/potassium (K) and yields (as well as profits) to update farmers' beliefs about the effectiveness of these two fertilizers. This treatment is similar to the intervention in Hanna, Mullainathan, and Schwartzstein (2014), where they presented farmers with a summary of the trials' findings that pod

[17] The application is also friendly to non-smart-phone users. There is a function called "send a message about the testing results and recommendations". A non-smart-phone user can use another person's smart phone to get the test result, and then send the test result and recommendations to her own phone as a message.

[18] During the baseline survey, our enumerators

size is important, and found that it induced farmers' learning without providing new data.

4) Control group: 300 farming households were untreated.

We list the timeline of the data collection activities and the implementation of first-phase experiment (second-phase in gray) as follows,[19]

Oct 2019 - Nov 2019	Pilot survey
Mar 2020 (late) - April 2020	Baseline survey
- - -	Mobile application collecting individualized soil testing results
Mar 2020 (late) - April 2020	Phase 1 interventions: four arms + three interventions
May 2020 - Mid July 2020	Fertilizer purchase and application season
Late Sep 2020 - Oct 2020	Harvest season
Nov 2020 - Dec 2020	1^{st} Follow-up Survey
- - -	Leaf color chart collecting recalled data on past greenness
Nov 2020 - Dec 2020	Phase 2 intervention: leaf color chart distribution
May 2021 - Mid July 2021	Fertilizer purchase and application season
Jul 2021 - Aug 2021	2^{nd} Follow-up Survey
Late Sep 2021 - Oct 2021	Harvest season
Dec 2021 - Jan 2022	3^{rd} Follow-up Survey for extended project
Nov 2022 - Dec 2022	Planned 4^{th} Follow-up Survey for extended project

We randomly assigned the 200 villages in our sample to four different arms (T1, T2, T3 and Control) in late March 2020, when China's economy was fully reopened.[20] The data collection activities lasted for two years: we surveyed farmers in late March 2020 before the season of fertilizer application (baseline), in November 2020 after the harvest season (first follow-up survey), and in summer 2021 after the second season-year of fertilizer application. In the surveys, we collected information on (i) input and output including yields, profit, land area, and other variables; (ii) purchase and usage of different fertilizers during multiple stages of plant growth; (iii) testing results

[19] We mark in gray the activities for the second-phase experiment which will be introduced later.

[20] China's economy was gradually reopened in March 2020 after coronavirus shutdowns and recovered fully in late March. Our experimental site, Leiyang, had very few identified cases and was in the first batch of reopening. The number of existing confirmed cases went down to zero by 02/28/2020 and Leiyang city has recorded no cases from 02/28 to date.

on individual soil quality from the nearest testing plot, distance from the farmer's plot to the testing plot, fertilizer recommendations predicted by the soil testing, and the gaps in fertilizer use between farmers' actual practice and the recommended use; and (iv) farmers' beliefs about the returns to greenness, effectiveness of nitrogen (N)/phosphorus (P) and / potassium (K).

The first-phase experiment was conducted in April 2020. We did several things to ensure the implementation of the design. For the T3 group (App + in-person AEA's training), in December 2019, we organized two training sessions for all the agricultural extension agents (AEA) in Leiyang, Hunan province, preparing for incoming interventions in the T3 group. During the two weeks before the survey, we provided comprehensive training to the enumerators, who were all local college students and could speak local dialect. During the survey itself, we also allocated several independent monitor team members to randomly join the interview process for quality inspection.

2.4 Data and Balance Test

Table 1 shows basic summary statistics from the baseline survey. The first four columns report the means for T_1, T_2, T_3, and control farmers; the last three columns report the difference between T_1 and control, T_2 and control, and T_3 and control, respectively. Panel A on farm characteristics shows that in 2020 average yield was about 460-470 kilograms per mu, with revenue per unit of land of 1096-1120 RMB. On average, for each mu of rice, farmers applied 36 kilograms N-P-K compound fertilizers and 20 kilograms nitrogen fertilizers. Turning to top-dressing phosphorus and potassium fertilizers, the adoption rate was quite low; most farmers (97% for phosphorus and 91% for potassium) didn't apply any top-dressing phosphorus or potassium during the growing stage. As a result, the application intensity for phosphorus and potassium, on average, is only 0.84/2.16 kilograms per mu.

Panel B presents soil testing results and predicted fertilizer recommendations. The vast majority of farmers can be linked to a tested plot within 0.2 kilometers (> 50%). One-third of the households had a distance lower than 0.1 kilometers. Panel C demonstrates farmers' perception of the return to greenness, beliefs about the effectiveness of potassium fertilizers,[21] and the attrition. We present the two related questions as follows,

To measure farmers' beliefs about the return to greenness:

> *What is the relationship between greenness and yield?*
> *1) The greener the leaves are, the better the yield is;*
> *2) No strong relationship;*
> *3) Inverted U-shape, yield increases first as greenness increases, and then decreases when greenness passes a certain threshold.*

Among these options, option 3) is well-documented in scientific and agronomic studies. Panel C shows that only a small proportion (7%) of farmers gave a correct response to the relationship between greenness and yields.

[21] Unfortunately, we only had baseline data for farmers' beliefs about the effectiveness of potassium, but no baseline data for farmers' beliefs about the effectiveness of phosphorus fertilizers.

To measure farmers' beliefs about the effectiveness of potassium, we ask the following question. We mark in bold the correct answer.

> *Which of the following micronutrients affects grains' density?*
> *(1=N, 2=P,* ***3=K****, 4=don't know it)*

We find that most of the farmers (98%) cannot give a correct response to this question, which is really puzzling.

Overall, there are no systematically significant differences between the treatment arms and control farmers in any of the variables, except for two that are statistically significant at the 10% level. The first is the recommended N-P-K compound fertilizer between the T_3 group ($App+Visit$) and the control group. The second imbalance appears in the attrition rate between T_2 and control farmers. Our attrition overall is quite low; roughly 3 to 8 farmers per treatment arm opt out of our study. This slight imbalance should not cause any major concerns. These between-group comparisons confirm the validity of our randomization.

3 First-phase Experiment Results

In this section, we establish the stylized facts that farmers simultaneously overuse nitrogen and underuse phosphorus/potassium and that such sub-optimal input choices can be corrected by cost-effective interventions. We show that the provision of customized fertilizer recommendations through a mobile application made farmers better off, and induced a reduction in nitrogen use and an increase in phosphorus/potassium use. In the next section, we study mechanisms under misspecified learning and design a theory-based intervention to test model predictions.

3.1 Actual Use Deviates from Recommended Use

We begin the analysis with graphical evidence that highlights some important features in the data. Figure 4 presents the number of households (y-axis) against the percentage of deviation (x-axis) between farmers' actual use and the recommended use of different fertilizers.[22] The green bar and red bar highlight the simultaneous overuse of nitrogen and underuse of phosphorus and potassium, respectively. The gaps in different fertilizer use are present for a large part of the domain, showing a 30-150% deviation in nitrogen overuse and 10-70% deviation in the underuse of phosphorus and potassium. The graphical evidence is consistent with the findings in Cui et al. (2018).

In Figures 10a, 11a, and 12a, we provide more detailed evidence on farmers' deviation in nitrogen fertilizer use in different cropping stages. Figure 11a shows that the mean application (blue line) overlaps with the mean recommendation (red line) in the planting stage, suggesting that nitrogen fertilizer use did not deviate from the recommendations in the planting stage. Figure 12a presents that the deviation of nitrogen use comes from farmers' practice in the growing stage (Used > Recommended). Turning to phosphorus fertilizers, we observe that the deviations (Used

[22] Deviation ratio = (Actual use - recommended use)/recommended use.

< Recommended) come from both the planting stage (Figure A6a in Appendix A) and the growing stage (Figure A7a in Appendix A). Such deviations are similar in the application of potassium fertilizers; we observe that the deviations (Used < Recommended) come from both the planting stage (Figure A9a in Appendix A) and the growing stage (Figure A10a in Appendix A).

3.2 Main Specification

To quantify the inefficiency from simultaneous overuse and underuse, we exploit the following regression specification:

$$\mathrm{Y}_{iv} = \beta_0 + \beta_1 * T1_v + \beta_2 * T2_v + \beta_3 * T3_v + \eta_{iv}$$

where i indexes farming households, v indexes villages, and Y_{iv} is the outcome variable of interest in the post-treatment period. It includes 1) The usage of different fertilizers; 2) Yield, profit, revenue, fertilizer cost and other input costs; 3) Farmers' valuation of the effectiveness of different fertilizers. T_1, T_2, and T_3 are indicator variables for different treatment arms, which equal one if the village is assigned to the only-soil-testing group, the mobile application group, and the App plus training group, respectively. For inference, we cluster standard errors at the village level to reduce any correlated errors within a village. Our coefficients of interest are β_1, β_2, and β_3, which measure the intention-to-treat effects of the three interventions.

3.3 Results

Table 2 presents the results for use of four different fertilizers (N, P, K, and N-P-K- compound fertilizers). In columns (1)-(3), we compute total nitrogen/phosphorus/potassium fertilizer use by converting compound fertilizer (N:P:K = 15:15:15) into individual fertilizers.[23] Start with column (1) where the outcome is the total nitrogen fertilizer usage (after conversion). While the mean intensity of nitrogen use in a control farming household was 30.84 kg/mu (equivalent to 462.6 kg per hectare), nitrogen use in the T_2 (mobile application) group dropped dramatically, by 3.92 kg/mu ($p < 0.05$), corresponding to a treatment effect of 12.17%. In addition, nitrogen use in the T3 group (mobile application + training visit) dropped by 4.42 kg/mu($p < 0.01$), corresponding to a treatment effect of 14.33%.

Turning to phosphorus, column (2) shows that, compared to an average of 14.74 kg/mu (equivalent to 221.1 kg per hectare) in control households, T2 and T3 interventions increased phosphorus usage by 2.34 kg/mu ($p < 0.01$) and 2.72 kg/mu ($p < 0.01$), respectively. Similarly, column (3) suggests that, compared to a control mean of 13.31 kg/mu (equivalent to 221.1 kg per hectare), T_2 and

[23] Total N = (Urea *46% + Compound fertilizer* 15%)/(46%).
Total P = (Calcium superphosphate * 39% + Compound fertilizer* 15%)/(39%).
Total K = (KCL *45% + Compound fertilizer* 15%)/(45%).
where Urea is the main nitrogen fertilizer widely used, containing 46% nitrogen. Calcium superphosphate is the main phosphorus fertilizer used, containing roughly 18%-20% P_2O_5, and hence 39% phosphorus. KCL is the main potassium fertilizer widely used, containing 45% potassium. The most widely used compound fertilizer contains 15% nitrogen, 15% phosphorus, and 15% potassium.

T_3 interventions increased potassium usage by 1.37 kg/mu ($p < 0.1$) and 2.89 kg/mu ($p < 0.01$). Not surprisingly, we do not detect any significant change in the use of nitrogen/phosphorus/potassium in the T_1 group, since it's hard for farmers to read and handle the information for soil analysis results.

The remaining columns demonstrates the treatment effects at the extensive and intensive margin across different cropping stages. Column (4) presents the treatment effects at the transplanting stage, showing evidence on the increase of compound fertilizer use in T_2 and T_3, which contributed to the reduction in total nitrogen in column (1) and the increase of total phosphorus/potassium in columns (2)/(3). Columns (5)-(9) present the treatment effects at the growing stages. While columns (6) and (8) report the growth in top-dressing phosphorus/potassium fertilizer use on average, columns (7) and (9) focus on the extensive margin — the proportion of households using phosphorus/potassium within each treatment arm. Although only 3% of farmers applied top-dressing phosphorus during the growing stage before the intervention, our T_2 and T_3 interventions substantially increased the share by 23.9 and 23.2 percentage points (both $p < 0.01$). Similarly, the treatment effects in T_2 and T_3 are 24.1 and 32.4 percentage points (both $p < 0.01$), compared to a control mean of 9% for the proportion of farmers using potassium fertilizers. As a whole, we conclude that Table 2 shows reduction in nitrogen at the intensive margin and an increase in phosphorus/potassium use at both the intensive and extensive margins.

Closing the yield gap. Table 3 explores the secondary outcomes due to farmers' re-optimization in fertilizer mix. Columns (1) and (2) show significant treatment effects on the yields. Addressing the overuse of nitrogen and underuse of phosphorus/potassium fertilizers led to a 22.74 and 31.65 kg/mu increase in yields in T_2 and T_3, compared to an average yield of 465.6 kg/mu in the control group. As shown in column (2), this treatment effect corresponds to 5.4% ($p < 0.1$) and 6.7% ($p < 0.05$) increases in yields. Column (3) shows that farmers in T_2 and T_3 had a higher profit than those in the control group, as a result of increasing revenues and unchanged costs. As shown in column (4), the total revenues of farmers in T_2 and T_3 went up by 68.57 Yuan/mu and 78.79 Yuan/mu, respectively, accounting for 6% ($p < 0.1$) and 6.9% ($p < 0.05$) growth in revenues. While column (5) shows that there was no significant change in fertilizer costs, column (6) suggests that the costs of other inputs, including labor, didn't go up.

In summary, we find that T_2 and T_3 both effectively helped farmers re-optimize fertilizer inputs and improve yields and profits.[24] These effects are not statistically different between T_2 and T_3.

3.4 How Large Was the Inefficiency?

We first compare our effects to other related interventions to increase/reduce fertilizer usage. First, in terms of nitrogen-fertilizer application, Chen et al. (2014) estimate that integrated soil–crop system management (ISSM)-based recommendations resulted in higher yields (18–35%) and a reduction in nitrogen usage (4–14%). Similarly, Cui et al. (2018) show that the rollout of the ISSM

[24] We also summarize these results in Figures 10b for nitrogen fertilizers, A5c for phosphorus fertilizers, A8c for potassium fertilizers and A11b for yields.

program in China induced a reduction in the use of nitrogen by 14.7–18.1%, an overall yield improvement by 10.8–11.5% and a reduction in greenhouse gas emissions by 4.6–13.2%. We find smaller effects in the reduction of nitrogen (14%) and yields (5%-7%), but our intervention appears to be less costly and is cost-effective. Second, with regard to other fertilizers, Duflo, Kremer, and Robinson (2011) suggest that offering free fertilizer delivery immediately increased the proportion of farmers using fertilizer by 33%, and using $\frac{1}{2}$ teaspoon of top-dressing fertilizers per hole increased farmers' income by 15%. Our intervention also increased the proportion of farmers using phosphorus/potassium by 19.1/22.5 percentage points, as well as increasing their revenue by roughly 7%. Our treatment effects lie between these studies, suggesting that policy-makers can find a unified solution to resolve the simultaneous overuse and underuse. We also quantify the efficiency loss from incorrectly using fertilizers by the profits that are lost from misuse. The estimate suggests that, if all 440 million mu of rice plots[25] experienced the same level of adoption, then total revenues of these lands could go up by 30 billion RMB without increasing costs – not to mention the benefits to other crops and the environment.

We also compare our impacts to the research on Information and Communications Technology (ICT) and agricultural development. Cole and Fernando (2021) estimate the return of a mobile-phone based agricultural advice service provided to farmers in India, suggesting that it increased yields in cumin by 28% and cotton by 8.6%. We find smaller effects of 5-7% yield increases, perhaps because our mobile application only focuses on the optimization of fertilizer application, while theirs also directly delivered time-sensitive information such as weather forecasts and pest planning strategies to farmers. Such results show that ICT can serve as both a complement to and a substitute for the traditional agricultural extension service. There is longstanding concern that traditional agricultural extension has not always provided the most useful information for farmers or given them knowledge about alternative farming practices (Beaman et al., 2021). This is especially the case for farmers in less developed countries, especially smallholders, who may be ill-informed of sustainable fertilizer practices if the agricultural extension support is not adequate. Our interventions could also address such inefficiency in agricultural extension service provision by using the smart mobile application, which apparently reduced the administrative costs and agents' travel costs, but led to a similar degree of yield growth (yield effects of T_2 and T_3 are quite close).

4 Theoretical Framework

A natural question is why farmers are simultaneously overusing and underusing different fertilizers. This misuse pattern is really puzzling because fertilizers have been used for a long time. In addition, they are divisible and thus suitable for experimentation. In this section, we explore the mechanisms behind such misuse of technologies.

[25] 1.8 billion mu arable land in total for different kinds of crops in China.

4.1 Stylized Facts and Survey Evidence

To answer this question, we dig deeply into the different functions of different fertilizers. Fertilizers have three vital dimensions —nitrogen (N), phosphorus (P), and potassium (K) — among which 1) nitrogen (N) produces salient signals by increasing the greenness of the crops during the growing stage, while 2) phosphorus (P) and 3) potassium (K) do not generate salient signals. Specifically, phosphorus (P) boosts the root length and changes the timing of flowering, while potassium (K) enhances the density of the rice grains, which cannot be easily observed.

Figure 5a presents the recommended practice for optimal fertilizer application at multiple cropping stages. The cropping cycle is divided into three stages, shown from left to right: 1) transplanting/planting stage, in which basal fertilizer application is required; 2) growing stage, in which top-dressing fertilizers are strongly recommended; and 3) ripening and harvest stage, in which no fertilizer should be applied. During the transplanting stage, farmers are advised to apply N-P-K compound fertilizers as basal fertilizers, while the extra top-dressing nitrogen and phosphorus/potassium fertilizers are recommended to be used during the growing stage. Meanwhile, in the growing stage, farmers notice the changes in the greenness of the rice plant.

Figure 5b shows farmers' actual practice in fertilizer application before the interventions with the number of households on the y-axis and the number of weeks after transplanting on the x-axis. The red bar indicates that, consistent with the recommended timeline, almost all 1,200 farmers applied N-P-K compound fertilizers as the basal fertilizers during the transplanting stages. However, the green, orange, and blue bars show that, during the growing stage, farmers' actual application of top-dressing fertilizers deviated from the recommendations. We observe that farmers applied and adjusted only nitrate fertilizer (N) during the growing stage (>= second week), but did not add phosphorus (P) and potassium (K). Combining with the findings presented in Table 2 that the overuse of nitrogen was mainly from the growing stage, we hypothesize that, because farmers can observe the greenness signals that reflect the effectiveness of nitrogen during the growing stages, they adjust nitrogen use accordingly. However, since they cannot receive any salient feedback from applying phosphorus (P) and potassium (K), they could not evaluate their effectiveness and didn't use them.

Furthermore, we elicit farmers' beliefs about the relationship between yield and greenness. We present the question as follows,

> *What is the relationship between greenness and yield?*
> *1) The greener the leaves are, the better the yield is;*
> *2) No strong relationship;*
> *3) Inverted U-shape, yield increases first as greenness increases, and then decreases when greenness passes a certain threshold.*

Figure 6 plots farmers' beliefs about the relationship between greenness and yields. Most farmers believed that the greenness is always positively associated with the yields (option 1). Only a small proportion of farmers (less than 7%) understood the true production function between greenness and outputs -— an inverted U-shape relationship (option 3), which is well documented in agricultural studies. Given this incorrect prior due to salient signals from nitrogen, farmers persistently over-applied nitrogen because of their overestimation of the return to greenness (misspecified production function).

Based on these stylized facts, we leverage recent behavioral economics theories on misspecified learning (Heidhues, Kőszegi, and Strack, 2021) to develop our conceptual framework. We outline farmers' decision-making problem as a farmer faces multiple technologies (different fertilizers in this context), in which their overestimation of one technology leads to distorted actions in adoption, and then such distorted actions in turn affect farmers' beliefs about the effectiveness of different technologies. We also demonstrate how farmers' actions and beliefs evolve over time and converge to a sub-optimal equilibrium.

4.2 Setup

The objective environment: In each period $t \in \{1, 2, 3, \ldots\}$, a representative farmer produces observable profit (output) $\pi_t \in \mathbb{R}$ according to the twice differentiable profit function $\Pi(a_t, b_t)$ which depends on her action $a_t \in (\underline{a}, \bar{a}) = A$ and an external state $b_t \in \mathbb{R}$ beyond her control. Similar to Hanna, Mullainathan, and Schwartzstein (2014), to capture the idea that farmers have to learn about two types of fertilizers at the same time — nitrogen (N) and phosphorus (P)/potassium (K) — her production function has two dimensions that contribute to profit:

$$\Pi(a_{Nt}, b_{Nt}, a_{Kt}, b_{Kt}) = f_1(a_{Nt}) \exp(b_{Nt}) - c_N a_{Nt} + f_2(a_{Kt}) \exp(b_{Kt}) - c_K a_{Kt}, \tag{1}$$

Where the functions f_1 and f_2 are concave, $f_i > 0, \lim_{a \to 0} f_i'(a) = \infty, \lim_{a \to \infty} f_i'(a) = 0$. $a_{Nt}, a_{Kt} > 0$ is the amount of N-fertilizer/K-fertilizer that farmers use at t, c_{Nt} and c_{Kt} are the normalized unit costs of different fertilizers, respectively. Without loss of generality, we assume that land area is fixed in a certain period and further do not include labor in this production function. The external states $\exp(b_{Nt})$ and $\exp(b_{Kt})$ are beyond her control, influencing the marginal product of a_{Nt} and a_{Kt}, and are correlated across farmers and time. For example, we can interpret $\exp(b_{Nt})$ and $\exp(b_{Kt})$ as the realized effectiveness of nitrogen (N) and phosphorus (P)/potassium (K).

Assume that

$$b_{Nt} = \Theta_N + \epsilon_{Nt}, b_{Kt} = \Theta_K + \epsilon_{Kt}$$

where $\Theta_N, \Theta_K \in \mathbb{R}$ are the underlying fixed fundamentals and ϵ_{Nt} and ϵ_{Kt} are independent normally distributed random variables with mean zero and variance σ_N^2, σ_K^2. To be specific, we can interpret Θ_N and Θ_K as **the average effectiveness** of nitrogen fertilizer and phosphorus/potassium fertilizers.

Farmers' subjective beliefs. Farmers' prior is that Θ_N and Θ_K are normally distributed with means θ_{N0} and $\tilde{\theta}_{K0}$, and with variances v_{N0} and v_{K0}, respectively. While the farmer understands the general form of profit function, she may misunderstand the specification of production function f:

$$\tilde{\Pi}\left(a_{Nt}, b_{Nt}, a_{Kt}, \tilde{b}_{Kt}\right) = \tilde{f}_1\left(a_{Nt}\right)\exp\left(b_{Nt}\right) - c_N a_{Nt} + f_2\left(a_{Kt}\right)\exp\left(\tilde{b}_{Kt}\right) - c_K a_{Kt}, \tag{2}$$

where $\tilde{f}_1(\cdot)$ indicates the misspecified production function of nitrogen fertilizers. $\tilde{b}_{Kt}$ captures farmers' subjective beliefs about the effectiveness of phosphorus and potassium fertilizers, which could measure their misunderstandings of the state variable (b_{Kt}).

More specifically and without loss of generality, we define $\tilde{f}_1(\cdot) = \lambda\tilde{f}_1(\cdot)$, and rewrite farmers' subjective beliefs about the profit maximization problem as:

$$\tilde{\Pi}\left(a_{Nt}, b_{Nt}, a_{Kt}, \tilde{b}_{Kt}\right) = \lambda f_1\left(a_{Nt}\right)\exp\left(b_{Nt}\right) - c_N a_{Nt} + f_2\left(a_{Kt}\right)\exp\left(\tilde{b}_{Kt}\right) - c_K a_{Kt}, \tag{3}$$

Where λ represents the degree of misspecification in the production function due to her misperception about the return to greenness. Given her model, the farmer updates her beliefs about the fundamental in a Bayesian way and chooses her action in each period to maximize her perceived discounted expected profits.

To equalize the misspecified profit function and true function, we impose some key assumptions.

Assumption 1): $\lambda > 1$. The farmer overestimates the return to greenness. This assumption is aligned with our survey evidence shown in Figure 6, i.e., farmers believe that greenness always leads to a better yield. One possible explanation is that fertilizer use has increased across time as shown in Figure 1. It is plausible that farmers would have formed the belief that more greenness brings greater yields since they have spent most of their time on the upward-sloping part of the curve. Such explanation corresponds to the mechanism of "what you see is all there is" (Enke, 2020).

Assumption 2): The farmer only misunderstands the state (b_{Kt}), her subjective belief $\tilde{b}_{Kt}$, regarding the effectiveness of P/K-fertilizers, since P/K-fertilizers don't produce salient signals during the growing stage. For the state b_{Nt} regarding the effectiveness of N-fertilizers, the farmer has no misunderstanding and will correctly learn it in a certain period since nitrogen produces salient signals and timely feedback. Farmers can learn its effectiveness by tracking the greenness. This assumption is also consistent with the survey evidence. In our sample, 93.8% of farmers understand that the greenness is enhanced by nitrogen (N), while only 4.4% and 2.5% of farmers could give the correct response regarding the effect of phosphorus (P) on the root development and the effect of potassium (K) on the grain's density.

Assumption 3): For any action $a_{Nt}\ a_{Kt} \in \mathbb{R}$, and any state $b_{Nt}, b_{Kt} \in \mathbb{R}$, there exists a farmer's subjective belief $\tilde{b}_{Kt}$ such that $\Pi(a_N, a_K, b_N, b_K) = \tilde{\Pi}(a_N, a_K, b_N, \tilde{b}_K)$. Compared to Heidhues, Kőszegi, and Strack (2021), we slightly modify this condition to guarantee that farmers can find an explanation for any output/profit she observes. Thus, Bayes' rule can specify beliefs.[26]

Updating rule of θ_N and $\tilde{\theta}_K$. The farmer chooses actions (a_N and a_K) in each period to

[26] We also need to impose some weak technical assumptions as Heidhues, Kőszegi, and Strack (2021) do. Please see Appendix B for reference.

maximize her perceived expected profits ($\tilde{Q}$) in that period. Furthermore, as the farmer's priors (Θ_N, Θ_K) are normally distributed with mean ($\theta_{N0}, \tilde{\theta}_{K0}$) and variance ($v_{N0}$, v_{K0}), and her beliefs b_N and $\tilde{b}_K$ are independent and identically distributed (i.i.d.) based on $\mathcal{N}\left(\Theta, \sigma_N^2\right)$ and $\mathcal{N}\left(\Theta, \sigma_K^2\right)$, respectively, the Bayes' updating rules are presented as follows.

Updating rule of θ_N. At the end of period $t \geq 1$, her posterior belief is that Θ_N is normally distributed with mean:

$$\theta_{Nt} = \frac{\frac{\sigma_N^2}{v_{N0}}\theta_{N0} + \Sigma_{s=1}^t b_{Ns}}{\frac{\sigma_N^2}{v_{N0}} + t}$$

and variance:

$$v_{Nt} = \frac{1}{v_{N0}^{-1} + t\sigma_N^{-2}}$$

Where θ_{N0} is the initial prior of the realized state b_N.

Updating Rule of $\tilde{\theta}_K$. At the end of period $t \geq 1$, her posterior belief about K is that $\tilde{\Theta}_K$ is normally distributed with mean:

$$\tilde{\theta}_{Kt} = \frac{\frac{\sigma_K^2}{v_{K0}}\tilde{\theta}_{K0} + \Sigma_{s=1}^t \tilde{b}_{Ks}}{\frac{\sigma_K^2}{v_{K0}} + t}$$

and variance:

$$v_{Kt} = \frac{1}{v_{K0}^{-1} + t\sigma_K^{-2}}$$

In terms of phosphorus (P)/potassium (K), since there is a misspecification in the profit function involved with b_K, the farmer updates her belief about the effectiveness of phosphorus (P)/potassium (K) based on the misspecified subjective belief $\tilde{b}_{Kt}$.

4.3 Predictions and Simulations:

After observing profit π_t contributed by the realized states b_{Nt} and b_{Kt}, the farmer believes that the realized state was b_{Nt} and $\tilde{b}_{Kt}$ satisfying:

$$\tilde{\Pi}\left(a_{Nt}, a_{Kt}, b_{Nt}, \tilde{b}_{Kt}\right) = \pi_t = \Pi\left(a_{Nt}, a_{Kt}, b_{Nt}, b_{Kt}\right) \tag{4}$$

Hence we can derive the distorted belief about the effectiveness of phosphorus/potassium $\tilde{b}_K$. See Appendix B - Model Proofs for more details.

We summarize and present the key predictions as follows,

If a farmer overestimates the return to greenness and hence has a misspecification in the production function ($\lambda > 1$), then:

Prediction 1): ${a'_N}^* > a_N^*$, $\tilde{a}_K^* < a_K^*$, and $\tilde{b}_K^* < b_K^*$, where ${a'_N}^*$ and $\tilde{a}_K^*$ are the optimal nitrogen and phosphorus (P)/potassium (K) usages under the misspecified model. a_N^* and a_K^* are the the optimal nitrogen and phosphorus (P)/potassium (K) applications under the true model.

And $\tilde{b}_K^*$ and b_K^* represent farmers' beliefs about the effectiveness of phosphorus and potassium fertilizers in the equilibrium under the misspcified and true models, respectively.

The misspecified model induces **persistent** overuse of nitrogen fertilizer and under-investment in phosphorus (P)/potassium (K), as well as undervaluation of the effectiveness of P/K. Thus, farmers are trapped in a sub-optimal equilibrium. They would be better off by re-optimizing the mix of different fertilizers.

Prediction 2): If $a_N \to a_N^*$, $\tilde{a_K} \to a_K^*$, then $\tilde{b}_K \to b_K$ if λ remains unchanged or decreases.

If farmers are nudged to take the correct actions, then their subjective beliefs $\tilde{b}_K$ about the effectiveness of P/K will move toward the true state. As such, the undervaluation of P/K will decrease. The derivations of prediction 2) are detailed in Appendix B.

Prediction 3): Correcting the overestimation/misspecification of the return to greenness leads to $\{a_N,\, a_K,\, \tilde{b}_K\}$ converging to the true values. Specifically, the correction could reduce farmers' nitrogen use to the optimal value immediately (one period), but also induce gradual learning on the utilization of phosphorus and potassium fertilizers.

Among these predictions, prediction (1) is examined by the results from Tables 2 and Table 3. We also present the simulations in Figure 7, showing the convergence of g function, θ_N, and $\tilde{\theta}_K$. Consistent with the theoretical predictions and survey evidence, Figure 7c suggests that farmers' beliefs about the effectiveness of P/K converges to an undervaluation equilibrium in a slow speed, while Figure 7b shows that their valuation of nitrogen converges faster.

5 Test of Predictions and Second-phase Experiment

In this section, we 1) run additional regressions to test model prediction (2) that farmers will place more value on phosphorus and potassium if their actions are corrected; 2) design and implement a second-phase experiment to test model prediction (3); and 3) discuss the results from the second-phase experiment.

5.1 Test of Predictions 2

Prediction 1) is tested in Tables 2 and 3, while, for prediction 2), we present the results in Table 4 using regression specification (1) in Section 3.2 based on first-phase interventions. Starting with column (1), the outcome variable is a binary dummy variable equal to 1 if farmers understand the relationship between greenness and yield correctly (inverted U-shape). Farmers in the T_2 and T_3 groups in the post-treatment period show highly improved understanding of this relationship. While only 6% of farmers gave the right response to this question in the control group, T_2 and T_3 interventions increased this proportion by 15 ($p < 0.01$) and 31 percentage points ($p < 0.01$). Therefore, the impact in T_2 solely reflects the effect of individual learning. The effect in T3, double that in T_2, suggests the presence of both individual learning and social learning due to farmers' interactions with the agricultural extension specialists.

For columns (2)-(5), we proxy farmers' beliefs about the effectiveness of different fertilizers with farmers' responses to the following questions. We mark in bold the correct response to each question.

a. Which of the following micronutrients is the main determinant for crop's greenness?
*(**1=N**, 2=P, 3=K, 4=don't know it)*
b. Which of the following micronutrients is the main determinant for the timing of flowering?
*(1=N, **2=P**, 3=K, 4=don't know it)*
c. Which of the following micronutrients is the main determinant for root length?
*(1=N, **2=P**, 3=K, 4=don't know it)*
d. Which of the following micronutrients is the main determinant for grains' density?
*(1=N, 2=P, **3=K**, 4=don't know it)*

In Table 4, the outcome variables in columns (2)-(5) are binary variables equal to 1 if farmers correctly answered the corresponding questions. Column (2) shows that farmers' understanding of the effectiveness of nitrogen on greenness in the T_1, T_2, and T_3 villages is not significantly different from that of the control group, since 95% of farmers already had correct beliefs. Column (3) presents the treatment effects of farmers' learning about the effects of phosphorus on flowering timing during the growing stage. Compared to the control mean of 13 percentage points, farmers in T_2 and T_3 improved their understanding of the effectiveness of phosphorus by 22 percentage points ($p < 0.01$) and 34 percentage points ($p < 0.01$). Similarly, column (4) reports the treatment effects of farmers' learning about the effects of phosphorus on root length, which shows the same pattern as column (3). Column (5) focuses on farmers' understanding of the effectiveness of potassium. Again, our interventions in T_2 and T_3 increased farmers' valuation of potassium by 17.1 and 36.8 percentage points, compared to 2.8% of farmers who understood the effect of potassium in the control group.

In summary, Table 4 provides direct evidence on model prediction (2) that farmers' undervaluation of the effectiveness of phosphorus/potassium decreases if their distorted actions are corrected, and the valuation of the effectiveness of nitrogen remains unchanged. Furthermore, we find a large difference in the treatment effects between T_2 and T_3, which suggests the effects of social learning. In T2 group, farmers figured out the effectiveness by themselves (individual learning). In the T_3 group, in addition to individual learning, farmers also had interactions with the agricultural extension agent, which enhanced their learning (social learning).[27]

5.2 Second-phase Intervention

To test the prediction 3) in Section 4, in this section, we design and implement a second-phase experiment, aiming to correct farmers' overestimation of return to greenness (changing λ to 1).

[27] Prior to our intervention, only 20% of farmers had ever been ever contacted by agricultural extension agents.

Leaf Color Chart (LCC). The leaf color chart, developed by the International Rice Research Institute (IRRI), contains six panels of different greenness, which match the green colors to the reflecting spectral signature of rice leaves. It is a cost-effective tool (4-8 USD) for real-time greenness/nitrogen (N) management, which can facilitate farmers' learning on the optimal greenness.[28] According to the IRRI's simulation and recommendation, the optimal greenness in two main growing stages — peak tiller formation (20-30 days after the transplanting) and spike differentiation (40-50 days after the transplanting) — should be around panel 3 or slightly above panel 3.[29] We also prepared an instructive brochure, showing farmers the correct timing, position, and method to use leaf color charts. Our enumerators stayed with farmers in person and, on average, spent 30 minutes demonstrating the correct usage of leaf color charts.

As shown in Figure 9, to enhance farmers' learning, we distributed the leaf color charts across the 150 villages that were selected into the first-phase interventions.[30] In Table 5, we conduct an additional balance test between farmers in $T1$ group and the control group, and do not find any significant difference between these two groups in the second-phase intervention. The timeline of the data collection activities and the implementation of second-phase experiment are listed as follows.

Oct 2019 - Nov 2019	Pilot survey
Mar 2020 (late) - April 2020	Baseline survey
- -	Mobile application collecting individualized soil testing results
Mar 2020 (late) - April 2020	Phase 1 interventions: four arms+ three interventions
May 2020 - Mid July 2020	Fertilizer purchase and application season
Late Sep 2020 - Oct 2020	Harvest season
Nov 2020 - Dec 2020	1^{st} Follow-up Survey
- -	Leaf color chart collecting recalled data on past greenness
Nov 2020 - Dec 2020	Phase 2 intervention: leaf color chart distribution
May 2021 - Mid July 2021	Fertilizer purchase and application season
Jul 2021 - Aug 2021	2^{nd} Follow-up Survey
Late Sep 2021 - Oct 2021	Harvest season
Dec 2021 - Jan 2022	3^{rd} Follow-up Survey for extended project
Nov 2022 - Dec 2022	Planned 4^{th} Follow-up Survey for extended project

[28]Similarly, Islam and Beg (2021) conducted a field experiment in Bangladesh, and found that leaf color chart intervention reduced nitrogen fertilizer use by 8% and increased yields by 7%. According to a back-of-the-envelope estimate, the cost-benefit ratio is about 1:9.

[29]See more details in Appendix C, which demonstrates the full guidance and application process.

[30]There are two concerns that prevent us from cross-randomization. First, our partners, the local government, hoped we could keep the experimental structure as simple as possible. Second, because we only have 50 villages in each treatment arm and the regressions are clustered at the village level, such design could help us keep statistical power when exploiting the combined effects of LLC, LCC + App, and LCC + App + Training Visit.

The second-phase intervention was implemented in December 2020, right after the first follow-up survey round in November 2020. In August 2021, we conducted a second follow-up survey round to explore the treatment effects of the leaf color chart intervention on fertilizer usage and farmers' beliefs about the effectiveness of different fertilizers. If our prediction (3) is valid, then we should be able to see the leaf color chart effectively correcting farmers' overestimation of the return to greenness, and simultaneously accelerating their learning about phosphorus and potassium.

5.3 Results from the Second-phase Intervention

Table 6 reports the regression results using the regression specification in Section 3.2. Focusing on farmers in T_1, the coefficient in column (1) shows that aggregate nitrogen usage immediately dropped by 3.76 kg per mu ($p < 0.05$), which is consistent with our model prediction and simulation showing that the learning about nitrogen is fast and substantial. As for phosphorus and potassium fertilizers, the positive coefficients in columns (2) and (3) for T_1, though not significant, suggest that, on average, the leaf color chart intervention may increase the aggregate use of phosphorus and potassium. The insignificance may be driven by the large variation in the outcome variables.

To explores the details of the effect of the leaf color chart on the use of phosphorus and potassium fertilizers, we decompose fertilizer applications into two stages: the planting stage and growing stage. Column (4) reports the N-P-K compound fertilizer use in the planting stage. Because leaf color charts are only effective during the growing stage, we do not observe a significant change for compound fertilizer use during planting stage. However, we indeed find inspiring results in columns (7) and (9), which report the proportion of farmers using phosphorus and potassium. The positive and significant coefficients in columns (7) and (9) suggest that, in the growing stage, the proportion of farmers using phosphorus and potassium increased by 6 percentage points ($p < 0.01$ and $p < 0.1$).

As suggested in Hanna, Mullainathan, and Schwartzstein (2014), agents' attention is limited, while the dimensions of technology are large. Consequently, an agent can only learn about the dimensions that she attends to. Similarly in our context, our treatment effects indicate that some farmers were starting to conduct their own experimentation with phosphorus and potassium fertilizers, perhaps because their attention was free from nitrogen due to the help of the leaf color charts and thus they were getting more confident about conducting experiments on phosphorus and potassium. Combining the results in columns (1), (2), (3), (7) and (9), we find that, consistent with model prediction (3), correcting farmers' overestimation of the return to greenness using leaf color charts not only instantly improved their learning on nitrogen, but also simultaneously accelerated their learning about phosphorus and potassium, although at a slower speed.

Turning to farmers in the T_2 group (App + leaf color chart) and T_3 group (App + AEA's training + leaf color chart), columns (1), (2), and (3) show that leaf color charts did not significantly enhance the existing treatment effects after the first-phase experiment on the aggregate usage of nitrogen, phosphorus and potassium, since many farmers in these two groups had already learned to re-optimize the usage of different fertilizers because of the first-phase experiment. In the meantime, the leaf color charts did not affect the existing treatment effects from the first-phase experiment either in the planting stage (column (4)) or growing stage (columns (5)-(9)). Since the soil testing and customized fertilizer recommendations have already provided farmers with precise guidance, there is no reason for the leaf color chart intervention to have an add-on effect.

Table 7 presents the effects of leaf color charts on farmers' perceptions of the return to greenness and beliefs about the effectiveness of different fertilizers after the second-phase intervention. Column (1) reports farmers' beliefs about the relationship between greenness and yield, while column (2) reports farmers' understanding about the effect of nitrogen on greenness. Columns (3) and (4) present farmers' understanding of the effects of phosphorus on flowering timing and root length, and column (5) shows farmers' understanding of the effects of potassium on grain's density. Farmers in the T_2 and T_3 groups, consistent with the results in Table 4, changed the overestimation of return to greenness (column (1)), kept high understanding of the effectiveness of nitrogen, and better understood the effectiveness of phosphorus and potassium (columns (3)- (5)).

For farmers in the $T1$ group who only received the leaf color chart treatment, column (1) shows that their overestimation of the return to greenness declined. The proportion of farmers understanding the inverted U-shape relationship between greenness and yield increased by 27.1 percentage points ($p < 0.01$) due to the leaf color chart treatment. This treatment effect is slightly lower than but comparable to that of the first-phase experiment (31.0 percentage points). Again, while farmers had no change in their understanding of the effect of nitrogen on greenness, since 95% of them already knew that, they also didn't increase their understanding of the effectiveness of phosphorus and potassium in the current period. Because farmers had not yet observed the yield,[31] they could not update their beliefs and realize the effectiveness of phosphorus and potassium. We expect to see the belief updating process after farmers produce and observe the new yields and profits, which inspires our future follow-up survey and research design.

Overall, compared to customized fertilizer recommendations, we find the second-phase intervention — leaf color charts — can also effectively correct farmers' overestimation of the return to greenness and improve learning of nitrogen/phosphorus/potassium, but in a gradual process. All of these results are summarized in Figures 10, 11, 12, as well as in Figures A5, A6, A7, A8, A9 and A10 in Appendix A. Nitrogen fertilizers were overused in the growing stage only, while phosphorus and potassium fertilizers were underused throughout (both planting stage and growing stage). From a policy design perspective, the former is more instant but costly, while the latter is more easily scalable but much slower. Policy makers can balance this trade-off based on budget and time constraints.

5.4 IV Strategy: Deviation in Fertilizer Application and Yields

In this subsection, we recover the yield response to the deviations between the actual fertilizer application and the benchmark recommendations based on soil testing. We first present evidence that our first-phase and second-phase interventions significantly reduced the gap in fertilizer use

[31] Our third-round survey was conducted in 2021 August, while the harvest season was in October.

between actual and recommended. In Table 8, columns (1), (2) and (3) present the treatment effect of the first-phase interventions on the gap between the actual application and recommended use. The outcome variables in columns (1), (2), and (3) are the nitrogen use gap [used - recommended], the phosphorus use gap [used - recommended], and the potassium use gap [used - recommended], respectively. In the first phase, $T2$ and $T3$ reduced the nitrogen gap by 3.6 kg/mu and 4.7 kg/mu, corresponding to a 50% dip in the control mean gap. Similarly, the interventions effectively closed the phosphorus gap and potassium gap by roughly 50%. Columns (4), (5) and (6) present the treatment effect of the second-phase interventions on the gap between the actual application and recommended use. Again, we find very similar effects on closing the gaps in fertilizer application. Since the testing data were acquired through GPS tracking, in Table A1 in Appendix A we show that the gaps are not associated with the distance between a farmer's plot and the nearest soil testing point, in either the first phase or second phase.

We then estimate the causal impact of the loss in yields by using the following IV strategy.

(1) Sample Selection. We limit the regression sample to those who overuse nitrogen fertilizers and underuse phosphorus/potassium fertilizers so that the underlying relationships between fertilizer gaps and yields are clearly defined. The overuse of nitrogen and underuse of phosphorus/potassium directly lead to lower yields. In Table A4 and A5 in appendix A, we relax this restriction and the results are still consistent.

(2) Main Outcomes. We calculate the gaps of each type of fertilizer by aggregating different stages, as N_Gap, P_GaP, and K_Gap, where N_Gap=(total nitrogen used - recommended), P_Gap=(total phosphorus used - recommended), and K_Gap=(total potassium used - recommended).

(3) Identification. In Table 8, we show that T_2 (App) and T_3 (App + Training) interventions effectively reduce nitrogen/phosphorus/potassium gaps, suggesting that T_2 and T_3 are valid and plausible instruments for these fertilizer gaps. But this raises an underidentification issue: we have three endogenous variables (N_Gap, P_GaP, and K_Gap) but only two instruments (T_2 and T_3). To address the underidentification constraint, we construct a new variable, Gap^2, which measures the Euclidean distance of the actual fertilizer application and the recommendations:

$$Gap^2 = (N\_Gap)^2 + (P\_Gap)^2 + (K\_Gap)^2$$

Then we take the log of Gap^2 and instrument $logGap^2$ with T_2 and T_3 in the two-stage least squares (2SLS) regression.

Equation (5) characterizes the causal effect of T_2 and T_3 interventions on Gap^2.

$$logGap_i^2 = \alpha' + \beta_1' T_{2i} + \beta_2' T_{3i} + \epsilon_i' \tag{5}$$

While equation (6) captures the effects of deviation in fertilizer application on yield.

$$Yield_i = \lambda + \delta \widehat{logGap_i^2} + \epsilon_i'' \tag{6}$$

Where $Yield_i$ indexes the yields (or $logyield$) of farmer i in the second round of survey (after the first-phase interventions). Our coefficient of interest δ measures the effect of deviation in fertilizer application on yields.

Table 9 reports the results of regressions (5) and (6). Column (1) presents the coefficients β_1' and β_2' in the first-stage regression, consistently suggesting that $T2$ and $T3$ effectively reduced the natural log of Gap^2 by roughly one unit. The second-stage regression result in column (2) suggests that, if the natural log of Gap^2 increases by one unit, then the yields will decline by 336.39 kg/mu. In column (3), we replace yields with the natural log of yields and find that a ten percent drop in Gap^2 will boost yields by 0.77%. To give a more intuitive magnitude, our $T1$ and $T2$ intervention in the first phase reduced Gap^2 by 100% (column (1)), corresponding to a 7.7% (336.39 kg/mu) increase in yields, which is consistent with our reduced form regression results of 7% in Table 3. To check the validity of $T2$ and $T3$ as instrumental variables, in Table A2 we replicate the IV-2sls regression using the baseline data and employing equations (5) and (6). We do not find any significance either in the first-stage or second-stage regressions since $T2$ and $T3$ did not affect the fertilizer applications and yields in the baseline data, which further confirms the validity of the IVs.[32]

Back-of-the-envelope Estimation for the Second-phase Intervention. To predict the yield response to the leaf color charts, we take coefficients in Table 6, 3.76 from column (1), 0.48 from column (2), and 0.173 from column (3). Then we obtain ΔGap^2 as:

$$\Delta Gap^2 = (N - 3.76 - \bar{N})^2 + (P - 0.48 - \bar{P})^2 + (K - 0.13 - \bar{K})^2 (N - \bar{N})^2 + (P - \bar{P})^2 + (K - \bar{K})^2$$

A naive estimation suggests that the leaf color charts reduced the log of Gap^2 by 0.44 points, indicating a 44% decline in Gap^2. Such a decrease corresponds to a potential yield gain of 3.4% or 16 kg/mu. In Table 10, we additionally present the effects of deviation in fertilizer application on revenues, fertilizer costs, and other costs using regression equations (5) and (6). Column (2) suggests that a ten percent drop in Gap^2 will increase revenues by 9.82 RMB per mu.[33] By plugging in the value of 44% decline in the Gap^2 , we estimate that the revenues would be increased by 43.21 RMB per mu, corresponding to a 4% increase in revenues per unit of land without changing the costs. These results are consistent when we employ the full sample from the second follow-up survey.

In summary, based on a back-of-the-envelope estimation, our second-phase intervention that

[32] In Table A3, we conduct another type of IV regression. In columns (1), (2) and (3), we instrument the gap in nitrogen use [Used - Recommended], gap in phosphorus use, and gap in potassium use separately with the $T2$ and $T3$ indicators. We then estimate the effects of each gap on yields separately in these columns. The results are quite consistent with the IV regression results in Table 9.

[33] We do not find significant impacts of Gap^2 on fertilizer costs and other input costs, which are consistent with previous findings in Section 3.3.

provides leaf color charts led to an increase in yields by roughly 3.4% and revenues by 4%.

5.5 Additional Discussions

Though this paper focuses on farming, our finding that the misperception of the first technology could affect the valuation on the effectiveness of the second technology has more general real-life implications. For example, for health conditions in less-developed areas, there is a trade-off between the use of antibiotics and improvement of hygiene conditions. Since the effects of antibiotics are immediate and can easily be noticed, while the effects of improving hygiene conditions on health are less salient, antibiotics are often over-taken. In this section, we discuss some issues with identification and interpretation, as well as some potential alternative explanations for nitrogen overuse.

Measurement Error in Self-Reporting. A natural concern is that our main outcomes of interest are self-reported, which means measurement errors and experimenter demand effects may affect the results. First, we argue that farmers have no incentive to misreport the usage of different fertilizers and yields, since our interventions can influence usage of different fertilizers in different directions. Second, we mitigate this concern by collecting farmers' responses on three different question sets regarding fertilizer use in different survey modules: 1) total usage of different fertilizers; 2) total purchase of different fertilizers; 3) fertilizer application in different growing stages and aggregate level of multiple stages. We find that data in these questions are quite consistent, which makes misreporting issues unlikely.

Other Learning Models. The existing leading learning models cannot explain both overuse and underuse; for example, in the model of learning-by-doing, farmers should be able to learn the value of different technologies with a lot of experience and exposure to natural variation. In the model of learning through noticing (Hanna, Mullainathan, and Schwartzstein, 2014), the selective attention mechanism induces farmers to either pay attention to and correctly use one input dimension or ignore and underuse that input dimension. However, these models cannot predict simultaneous overuse and underuse. Likewise, the procrastination cited in Duflo, Kremer, and Robinson (2011) and social learning in Wolitzky (2018) can only partially explain the underuse.

Supply Side Actions. Another question is why sellers aren't stepping in to correct farmers' beliefs about the effectiveness of different fertilizers, especially phosphorus and potassium. To answer this question, we first present survey evidence that 1) farmers did not lack access to phosphorus and potassium, shown in Figure 13a; and 2) only a small proportion (2.43%) of them took the advice from fertilizer sellers seriously and followed the recommendations in Figure 13b. One plausible explanation is that sellers, unlike us, didn't have science-based knowledge on the soil quality of individual plots, and therefore just provided the average recommendation, which is less convincing. Besides, by correcting farmers' beliefs to induce them to use more phosphorus and potassium, they would lose profits if farmers used less nitrogen.

Furthermore, the majority of misapplication took place during the growing stage, when farmers added additional top-dressing nitrogen, where they could have instead (or also) bought phosphorus and potassium, which does not cost any more than their current regime (apply nitrogen only). Figures 13c and 13d also show that farmers' application decisions were not driven by learning from others or low quality of fertilizers. The former argument is consistent with the learning from others literature by Wolitzky (2018), which shows that input information is much more difficult than output data to learn about from neighbors. The latter argument is consistent with the fact that fertilizer quality was high in our experimental site.[34]

Price Difference and Budget Constraints. 1) Price Difference. Farmers may use nitrogen more since it has a relatively lower price than the other two fertilizers. We directly elicit farmers' decision-making in nitrogen usage. In Figure 14a, contradicting the low-price explanation, survey evidence shows that farmers applied a given amount of nitrogen fertilizer, not because of low price, but directly due to greenness signals (Figure 14b). 2) Budget constraints. Farmers use nitrogen more since it has a lower price than the other two fertilizers. Contradicting this explanation, we find farmers only deviated from nitrogen fertilizer use during the growing stages when they received signals from crops (Figures 10, 11, and 12), but still used compound fertilizers (more expensive) during the planting stages.

Based on the above arguments, we believe that these plausible alternative explanations are unlikely to affect our main results that the overestimation of return to greenness indeed induces undervaluation and underuse of phosphorus and potassium.

6 Conclusions

In this paper, we used a two-phase RCT to investigate simultaneous overuse and underuse of different fertilizers in China and to understand the mechanisms behind this sub-optimal equilibrium. Farmers underuse phosphorus/potassium due to overestimation of the return to greenness caused by nitrogen. In their subjective beliefs, lower-than-expected output is rationalized by lower-than-true perception of the effectiveness of phosphorus/potassium fertilizers. Our interventions, including provision of customized fertilizer recommendations through a mobile application and leaf color charts, effectively reduced nitrogen fertilizer uses, increased phosphorus/potassium fertilizer applications, and increased yields and profits. We find direct and indirect evidence on two essential mechanisms: 1) the salient feature of the technology caused overestimation of the return to the input, resulting in overuse, and 2) the overestimation of the return to the salient technology led to undervaluation and underuse of the less salient technologies. Our interventions improved farmers' learning about the effectiveness of different fertilizers.

We now discuss the external validity, scalability, and policy implications of these results. We begin with a cost-benefit analysis. Based on the parameters from Cui et al. (2018), a back-of-the-envelope estimate suggests that, if all rice farmers, for a total of 440 million mu of paddy field in China, could adopt either the individual soil testing plus mobile application that provides

[34] See the official announcement: http://www.hunan.gov.cn/hnszf/hnyw/bmdt/201403/t20140314_4808623.html. On average, qualified fertilizers accounted for 93% of the total fertilizers sold in 2014 based on an inspection of 226 batches of fertilizer samples.

customized recommendations or leaf color charts, then total nitrogen fertilizer usage would be reduced by 1.76 million tons (N_2O emission reduction), CO_2-equivalent emissions could be reduced by 37.4 million tons, and yield would rise by 11 million tons (22 billion Chinese Yuan, equivalent to 3.4 billion USD). In addition, the mobile application is applicable to up to 15 types of crops. Regarding the coverage of soil testing, on average, the testing cost is about 100-150 RMB per plot (equivalent to 15.6 USD - 23.4 USD). In terms of the mobile application, the one-time cost for the coding is roughly 30,000 RMB and the annual cost of server maintenance is roughly 10,000 RMB. With regard to leaf color charts, each chart costs only 8-30 RMB. If we apply the soil testing and mobile application intervention to every 4-5 mu of land, then the average annual benefit will be more than twice the estimated cost. If we apply the leaf color charts intervention to all rice farmers, then the benefits would be more than 10 times as large as the cost for an average farmer with 7.5 mu of plots.

Given these results, a natural question is why farmers overestimate the return to greenness. There are several possible answers. First, consistent with the argument of Enke (2020) that what you see is all there is, fertilizer use has increased across time (see Figure 1, it is plausible that people believe that more greenness is equivalent to greater yields, because they have spent most of their time on the upward-sloping part of the fertilizer application curve. Second, the path dependence may play an important role in nitrogen overuse. Many farmers just follow extension agents' advice. However, the extension agents recommended a fertilizer use at the level of production maximization to follow the national food security strategy in the 1990s. Under such a scenario, smallholders might only respond to the average return to fertilizer, rather than the marginal return, because they only choose input once a year. In this case, it becomes more difficult for them to reach the optimum because they cannot continuously vary the fertilizer input and observe the corresponding outcomes. This argument is aligned with the findings in Ito (2014) that, in the electricity market, consumers responded to average price rather than marginal or expected marginal price. Other remaining questions include 1) why don't farmers make their own experimentation and 2) how does the social network affect the overuse and underuse of different fertilizers? These questions deserve a systematic investigation.

Understanding the case of fertilizer misuse in China also has general implications for other parts of the world. For instance, underuse of fertilizers is prevalent in developing countries and overuse of nitrogen is common in developed countries. We hope these results can shed light on future interventions and scalable solutions in the fight against low productivity in agriculture and global greenhouse gas emissions.

References

Allcott, Hunt, and Dmitry Taubinsky. 2015. "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market." *American Economic Review*, 105(8): 2501–38.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K Khandelwal, and Eric Verhoogen. 2017. "Organizational barriers to technology adoption: Evidence from soccer-ball producers in Pakistan." *The Quarterly Journal of Economics*, 132(3): 1101–1164.

Beaman, Lori, Ariel BenYishay, Jeremy Magruder, and Ahmed Mushfiq Mobarak. 2021. "Can network theory-based targeting increase technology adoption?" *American Economic Review*, 111(6): 1918–43.

Berkouwer, Susanna B, and Joshua T Dean. 2019. "Credit and attention in the adoption of profitable energy efficient technologies in Kenya."

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does management matter? Evidence from India." *The Quarterly Journal of Economics*, 128(1): 1–51.

Casaburi, Lorenzo, Michael Kremer, and Ravindra Ramrattan. 2019. "Crony Capitalism, Collective Action, and ICT: Evidence from Kenyan Contract Farming." Working paper.

Casaburi, Lorenzo, Michael Kremer, Sendhil Mullainathan, and Ravindra Ramrattan. 2014. "Harnessing ICT to increase agricultural production: Evidence from Kenya." *Harvard University.*

Chen, Xinping, Zhenling Cui, Mingsheng Fan, Peter Vitousek, Ming Zhao, Wenqi Ma, Zhenlin Wang, Weijian Zhang, Xiaoyuan Yan, Jianchang Yang, et al. 2014. "Producing more grain with lower environmental costs." *Nature*, 514(7523): 486.

Cole, Shawn A, and A Nilesh Fernando. 2021. "'Mobile'izing Agricultural Advice Technology Adoption Diffusion and Sustainability." *The Economic Journal*, 131(633): 192–219.

Corral, Carolina, Xavier Giné, Aprajit Mahajan, and Enrique Seira. 2020. "Autonomy and Specificity in Agricultural Technology Adoption: Evidence from Mexico." National Bureau of Economic Research.

Cui, Zhenling, Hongyan Zhang, Xinping Chen, Chaochun Zhang, Wenqi Ma, Chengdong Huang, Weifeng Zhang, Guohua Mi, Yuxin Miao, Xiaolin Li, et al. 2018. "Pursuing sustainable productivity with millions of smallholder farmers." *Nature*, 555(7696): 363.

Currie, Janet, Wanchuan Lin, and Juanjuan Meng. 2014. "Addressing antibiotic abuse in China: An experimental audit study." *Journal of development economics*, 110: 39–51.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2008. "How high are rates of return to fertilizer? Evidence from field experiments in Kenya." *American economic review*, 98(2): 482–88.

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

Dupas, Pascaline. 2014. "Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment." *Econometrica*, 82(1): 197–228.

Enke, Benjamin. 2020. "What you see is all there is." *The Quarterly Journal of Economics*, 135(3): 1363–1398.

Fabregas, Raissa, Michael Kremer, and Frank Schilbach. 2019. "Realizing the potential of digital development: The case of agricultural advice." *Science*, 366(6471).

Fishman, Ram, Avinash Kishore, Yoav Rothler, Patrick Ward, S Jha, and R Singh. 2016. "Can information help reduce imbalanced application of fertilizers in India." *Experimental evidence from Bihar. International Food Policy Research Institute (IFPRI) Discussion Paper*, 1517.

Foster, Andrew D, and Mark R Rosenzweig. 2010. "Microeconomics of technology adoption." *Annu. Rev. Econ.*, 2(1): 395–424.

Fudenberg, Drew, Giacomo Lanzani, and Philipp Strack. 2021. "Limit Points of Endogenous Misspecified Learning." *Econometrica*, 89(3): 1065–1098.

Gilbert, Natasha. 2012. "One-third of our greenhouse gas emissions come from agriculture." *Nature News*.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. 2020. "Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution." National Bureau of Economic Research.

Hamory, Joan, Edward Miguel, Michael Walker, Michael Kremer, and Sarah Baird. 2021. "Twenty-year economic impacts of deworming." *Proceedings of the National Academy of Sciences*, 118(14).

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2014. "Learning through noticing: Theory and evidence from a field experiment." *The Quarterly Journal of Economics*, 129(3): 1311–1353.

Harou, AP, Malgosia Madajewicz, Chris Magomba, Hope Michelson, Kevin Tschirhart, Nyambi Amuri, and Cheryl Palm. 2018. "Can Information Improve Investment? Effects Of Site-Specific Soil Recommendations On Fertilizer Demand." *Working paper*.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack. 2018. "Unrealistic expectations and misguided learning." *Econometrica*, 86(4): 1159–1214.

Heidhues, Paul, Botond Kőszegi, and Philipp Strack. 2021. "Convergence in models of misspecified learning." *Theoretical Economics*, 16(1): 73–99.

Islam, Mahnaz, and Sabrin Beg. 2021. "Rule-of-Thumb Instructions to Improve Fertilizer Management: Experimental Evidence from Bangladesh." *Economic Development and Cultural Change*, 70(1): 000–000.

Ito, Koichiro. 2014. "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing." *American Economic Review*, 104(2): 537–63.

Jensen, Robert. 2007. "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector." *The quarterly journal of economics*, 122(3): 879–924.

Jensen, Robert, and Nolan Miller. 2017. "Information, demand and the growth of firms: Evidence from a natural experiment in India." *Wharton School Working paper*.

Karing, Anne. 2018. "Social signaling and childhood immunization: A field experiment in sierra leone." *University of California, Berkeley*, 2.

Miguel, Edward, and Michael Kremer. 2004. "Worms: identifying impacts on education and health in the presence of treatment externalities." *Econometrica*, 72(1): 159–217.

Murphy, David MA, Dries Roobroeck, David R Lee, and Janice Thies. 2020. "Underground Knowledge: Estimating the Impacts of Soil Information Transfers through Experimental Auctions." *American Journal of Agricultural Economics*, 102(5): 1468–1493.

Rahman, KM, and Dunfu Zhang. 2018. "Effects of fertilizer broadcasting on the excessive use of inorganic fertilizers and environmental sustainability." *Sustainability*, 10(3): 759.

Suri, Tavneet. 2011. "Selection and comparative advantage in technology adoption." *Econometrica*, 79(1): 159–209.

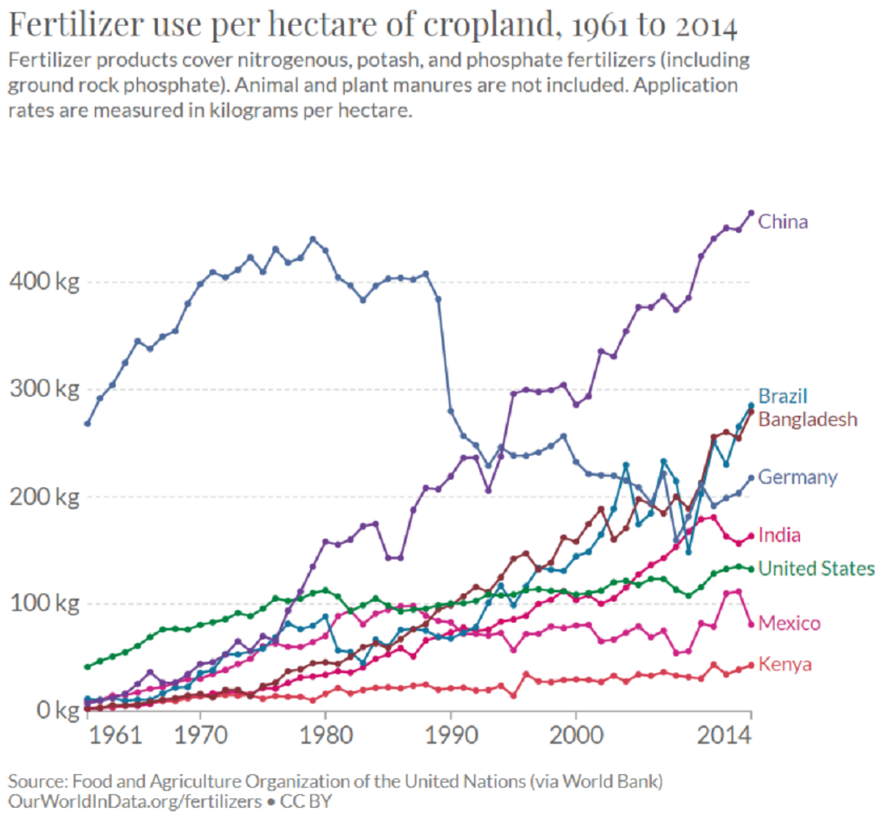
Tian, Hanqin, Rongting Xu, Josep G Canadell, Rona L Thompson, Wilfried Winiwarter, Parvadha Suntharalingam, Eric A Davidson, Philippe Ciais, Robert B Jackson, Greet Janssens-Maenhout, et al. 2020. "A comprehensive quantification of global nitrous oxide sources and sinks." *Nature*, 586(7828): 248–256.

Wolitzky, Alexander. 2018. "Learning from Others' Outcomes." *American Economic Review*, 108(10): 2763–2801.

Wu, Yiyun, Xican Xi, Xin Tang, Deming Luo, Baojing Gu, Shu Kee Lam, Peter M Vitousek, and Deli Chen. 2018. "Policy distortions, farm size, and the overuse of agricultural chemicals in China." *Proceedings of the National Academy of Sciences*, 115(27): 7010–7015.

Zhang, Xin, Eric A Davidson, Denise L Mauzerall, Timothy D Searchinger, Patrice Dumas, and Ye Shen. 2015. "Managing nitrogen for sustainable development." *Nature*, 528(7580): 51.

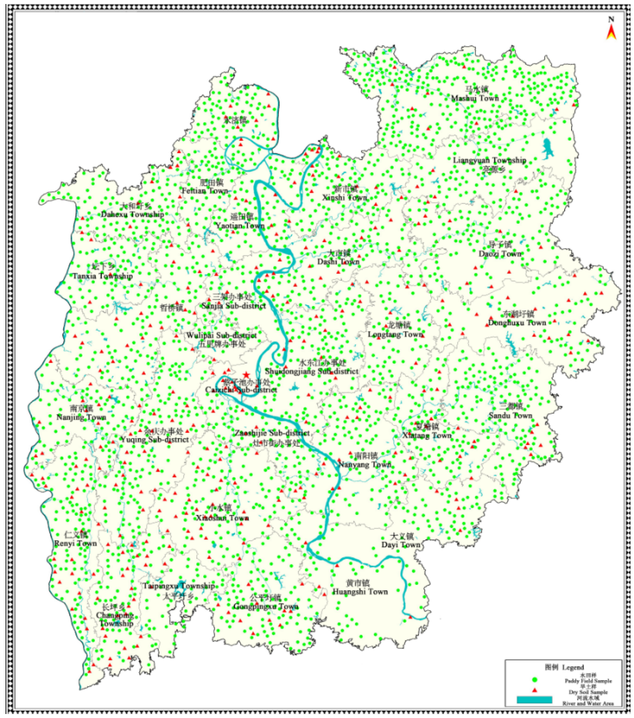
Figure 1: Cross-country Fertilizer Intensity, kg/ha



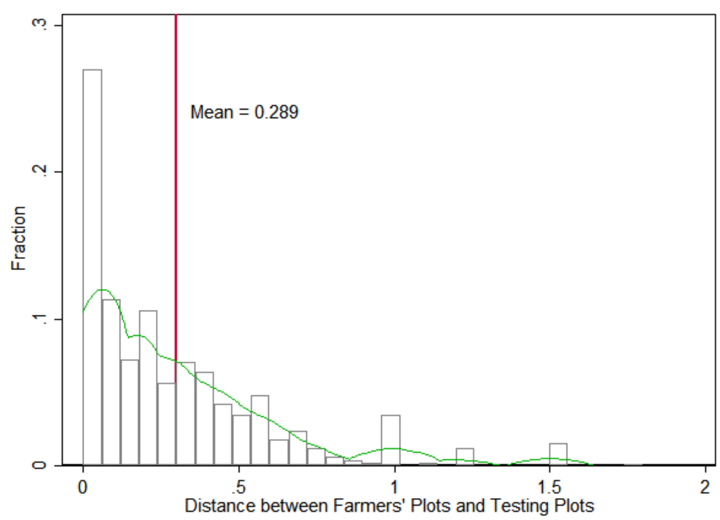
Note: This figure presents the pattern of cross-country fertilizer use per hectare from 1961 to 2014. The selected countries, in terms of fertilizer intensity in 2014, from top to bottom are China, Brazil, Bangladesh, Germany, India, United States, Mexico and Kenya. Most developing countries are on the upward-sloping part of the curve.

Figure 2: Link Farmers to the Universal Soil Analysis Points

(a) Universal Soil Testing Program in Leiyang, Hunan
(Green Dots as Testing Plots for Paddy Fields)

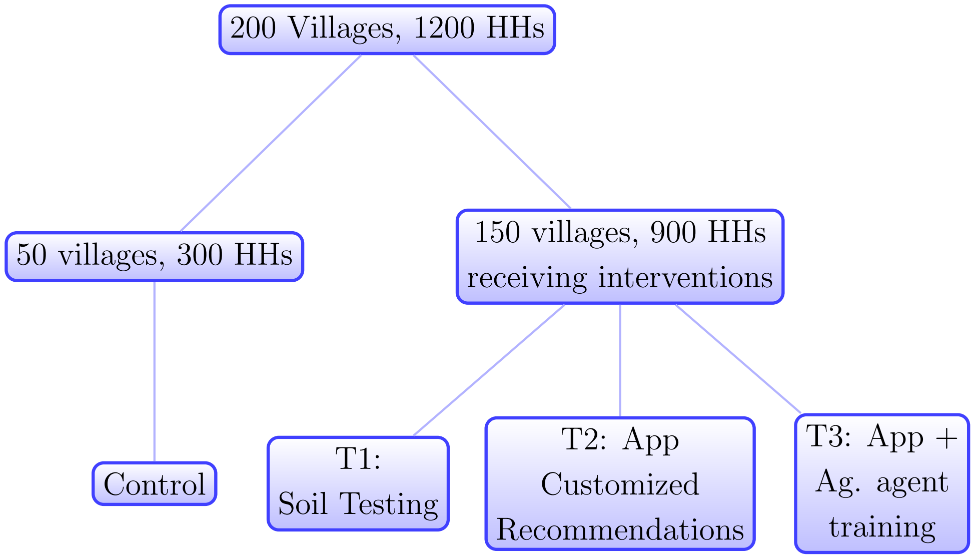


(b) Distance between Farmers' Plots and Nearest Testing Plots



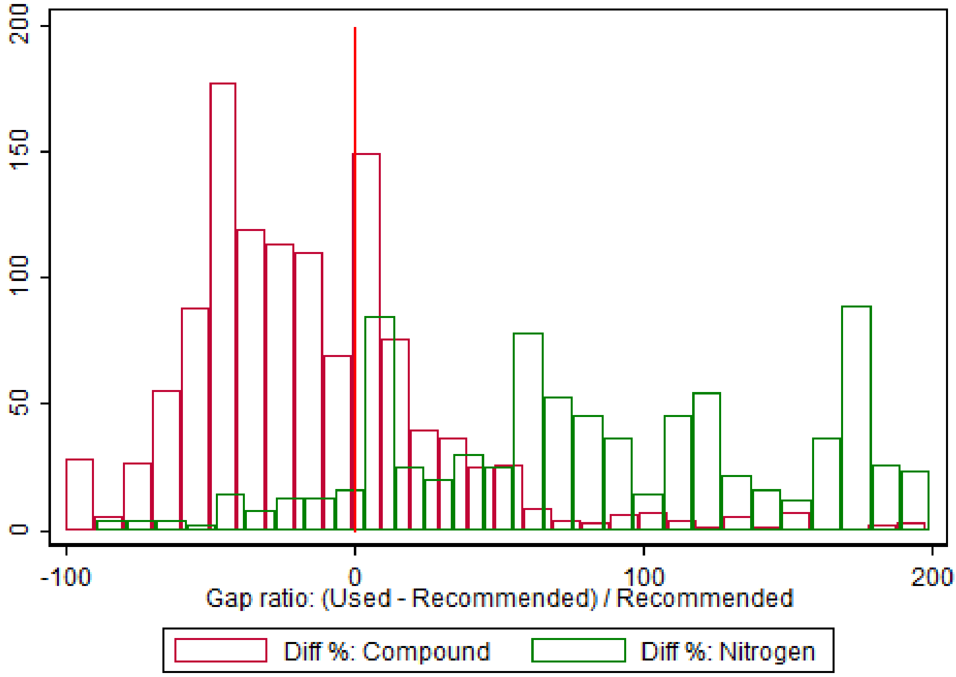
Note: This figure presents (a) the universal soil analysis in Leiyang, Hunan and (b) the distribution of distance between farmers' plots and the nearest testing point. In Figure 2a, each green dot is a testing point where a piece of soil sample was collected and analyzed in the lab for micronutrients component by agricultural extension specialists. We linked each farmer in the survey to the nearest testing plot (green dot), collected individual soil analysis results, and then generated customized fertilizer recommendations. Figure 2b presents the distribution of distance between a farmer's plot and the nearest soil testing point. On average, the distance is around 0.289, while the majority of distances were no more than 0.2km, and near 1/3 of household had a distance lower than 0.1 km.

Figure 3: First-phase experimental Design



Note: The figure presents the design and randomization for the first-phase experiment on 1,200 households in 200 villages. We randomly assigned 200 villages into four arms. In T1 group, farmers were only provided with soil analysis information (how many micronutrients in their plots). In T2 group, farmers received access to and training of our mobile application, which can not only display the soil analysis results, but also offer customized dynamic-fertilizer-application recommendations based on soil analysis. In T3 group, in addition to receiving the mobile application, farmers were also provided with a training session by agricultural extension agents for showing the experimental evidence on the relationship between phosphorus/potassium and yields, to increase their understandings on the effectiveness of the phosphorus/potassium fertilizers. The first-phase experiment was conducted in April 2020, which is before the season for purchase and application of fertilizers.

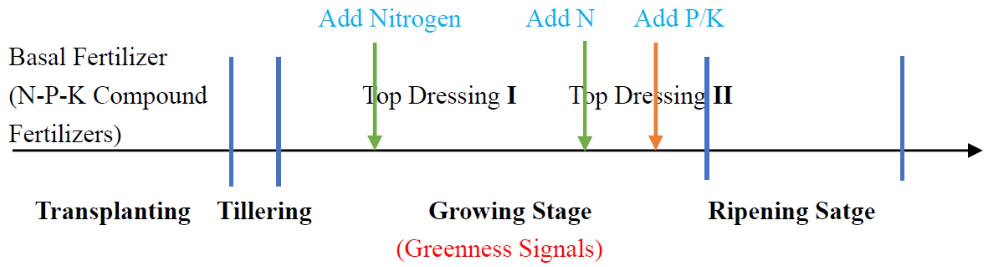
Figure 4: Gap in Fertilizer Application: (Used- Recommended)/Recommended



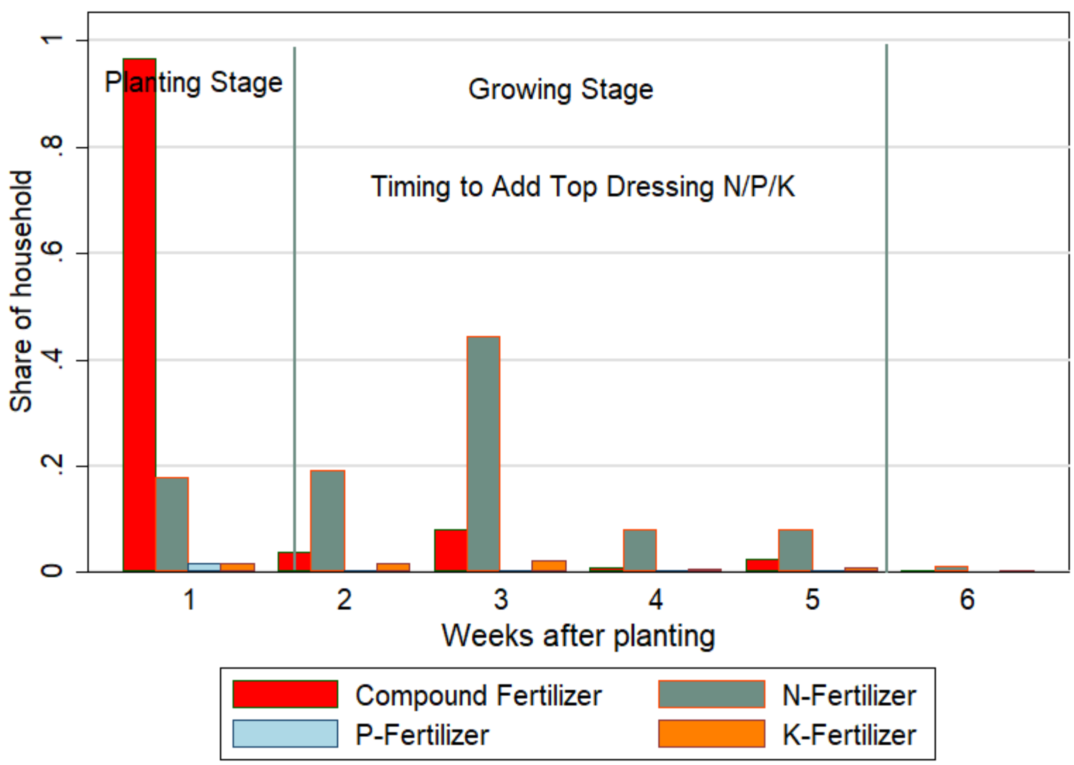
Note: This figure demonstrates the Gap in different fertilizer use between the actual use and the recommended use based on soil analysis data. The y-axis indexes the number of households, and the x-axis indexes the deviation percentage points, which equal to (actual Use - recommended)/recommended. The green bar and red bar show the deviation percentages for nitrogen and compound fertilizers, respectively.

Figure 5: Actual Application Deviates the Recommended Practice in Growing Stage

(a) Recommended Practice by Time



(b) Actual Practice by Time



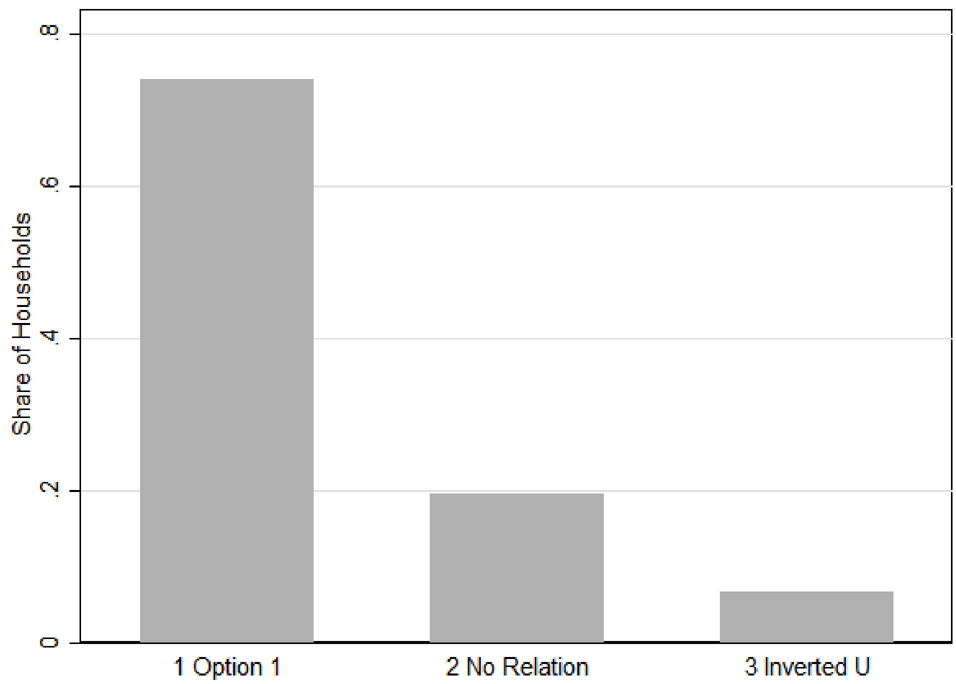
Note: These figures present how the actual fertilizer use from survey deviates the recommended practice. Figure 5a shows the recommended fertilizer application, divided into two main stages. During the transplanting stage, farmers are recommended to apply N-P-K compound fertilizer as the basal fertilizer, while during the growing stage, top-dressing nitrogen and phosphorus/potassium fertilizers are suggested to be used. In Figure 5b, the y-axis indexes the proportion of household applying different fertilizers, and the x-axis indexes different time. The left panel of Figure 5b is consistent with the recommended practice as shown in Figure 5a that almost all of the farmers were applying N-P-K compound fertilizer in the planting stage. However, the middle panel of Figure 5b shows that farmers only adjusted nitrogen use, but never added any phosphorus and potassium, which deviates the recommended practice.

Figure 6: Overestimation of the Return to Greenness

Option 1: The greener, the higher the yields are

Option 2: No strong Relationship

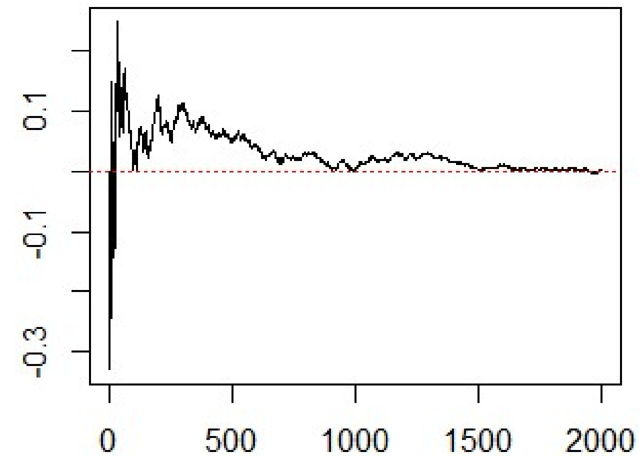
Option 3: Inverted-U Shape [✓]



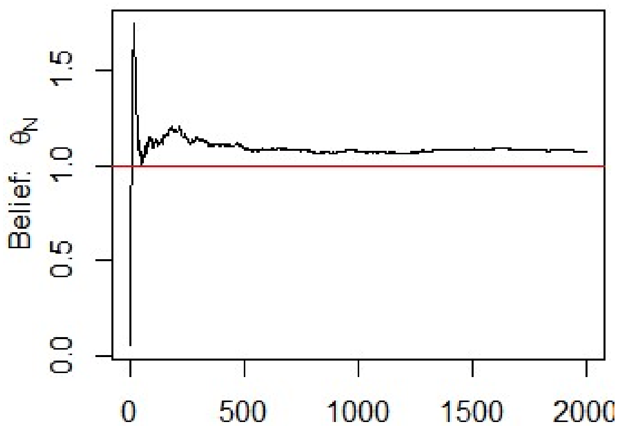
Note: This figure presents the farmers' beliefs about the relationship between greenness and yields. The y-axis indexes the proportion of households, while x-axis lists three different options. We can find that most of survey farmers, believed the greener the leaves are, the higher the yields are, suggesting an overestimation of the return to greenness.

Figure 7: Convergence of Beliefs about the Effectiveness of Different Fertilizers

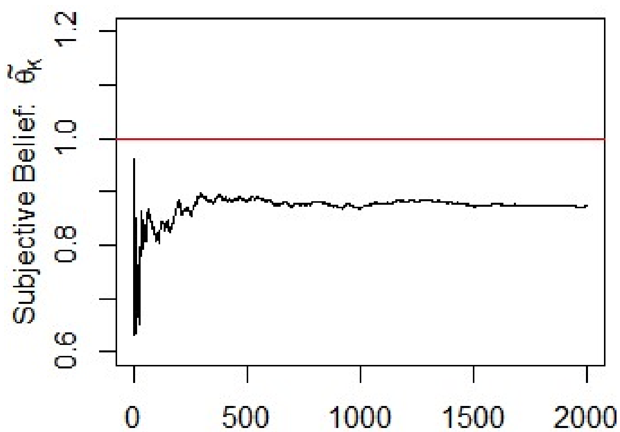
(a) Convergence of g Function ($g = 0 \Rightarrow$Equilibrium)



(b) Learning Effectiveness of Nitrogen

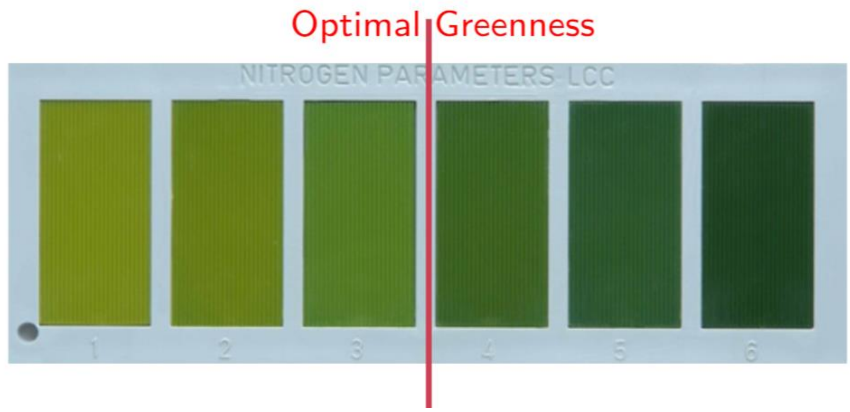


(c) Learning Effectiveness of Phosphorus/Potassium



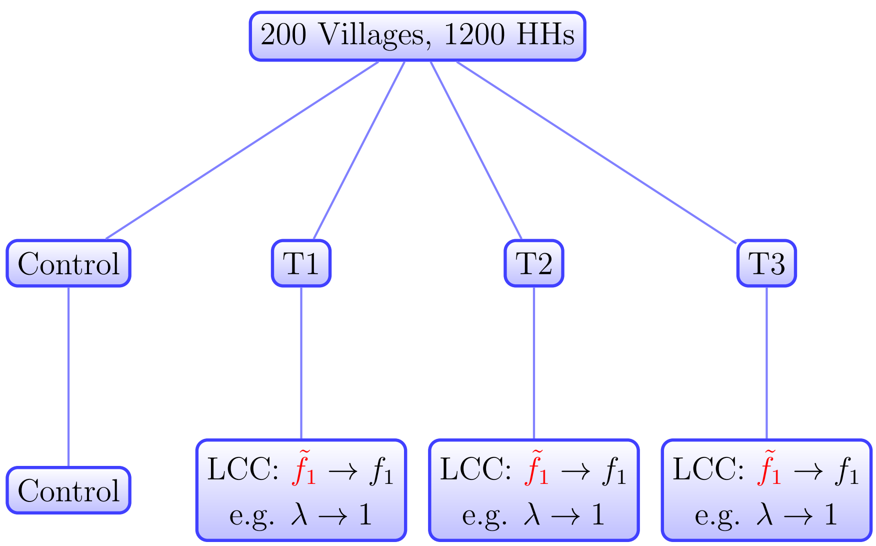
Note: This figure presents simulations about farmers' Bayesian learning on the effectiveness of nitrogen- v.s. phosphorus/potassium fertilizers. The top panel shows the convergence of g function, indicating the new equilibrium. Figure 7b displays farmers' belief about the effectiveness of nitrogen fertilizer compared to a true value of 1. Figure 7c depicts farmers' belief about the effectiveness of phosphorus/potassium fertilizer relative to a true value of 1, which is slower and undervalued.

Figure 8: A Sample of Leaf Color Chart



Note: The figure presents the six-color leaf color chart and the suggested optimal greenness.

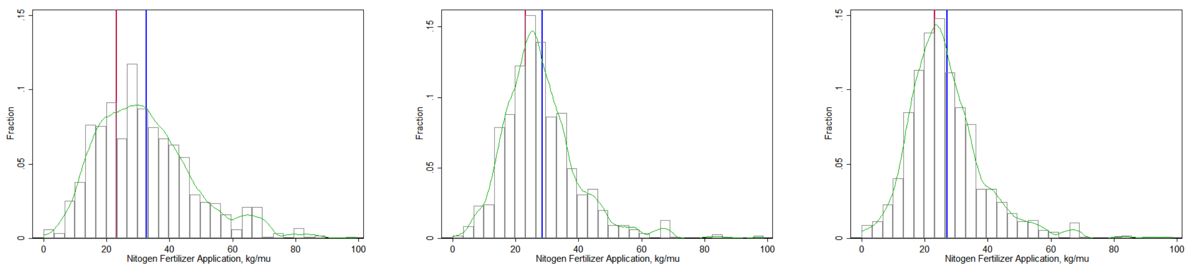
Figure 9: Second-phase Experimental Design: Correcting Overestimation



Note: The figure presents the design and randomization for the second-phase experiment on 1,200 households in 200 villages. Farmers in the existing treatment groups, T1, T2, and T3, all received the leaf color chart intervention. The target is to test the model prediction (3) by changing farmers' overestimation on the return to greenness and comparing the effects on T1 and control. The intervention was conducted in December 2020, after the harvest season and before the next fertilizer application season.

Figure 10: Total Nitrogen Application [Used >> Recommended]

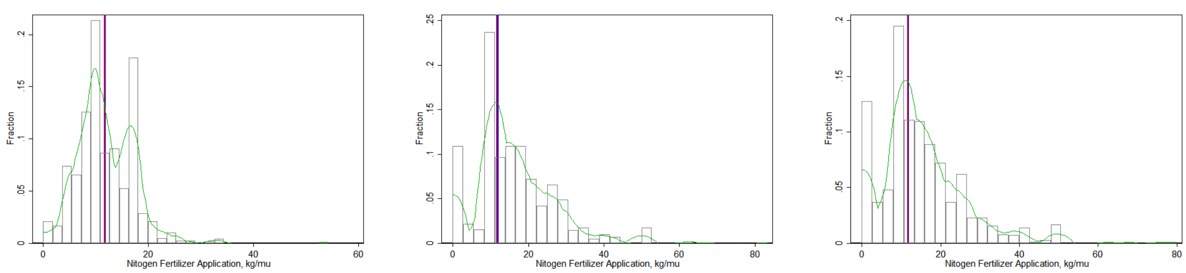
(a) Before Interventions, 2019 (b) After Phase 1 Intervention, 2020 (c) After Phase 2 Intervention, 2021



Note: This figure shows the distribution of total nitrogen fertilizer application (kg/mu) in the baseline (2019 season), after the first-phase interventions (2020 season), and after the second-phase intervention (2021 season). The red line indicates the mean of nitrogen recommendations, while the blue line shows the mean of actual use. The figure shows a clear pattern that the deviation in total nitrogen application decreases after our interventions.

Figure 11: Nitrogen Application in the Planting Stage [Used ≈ Recommended]

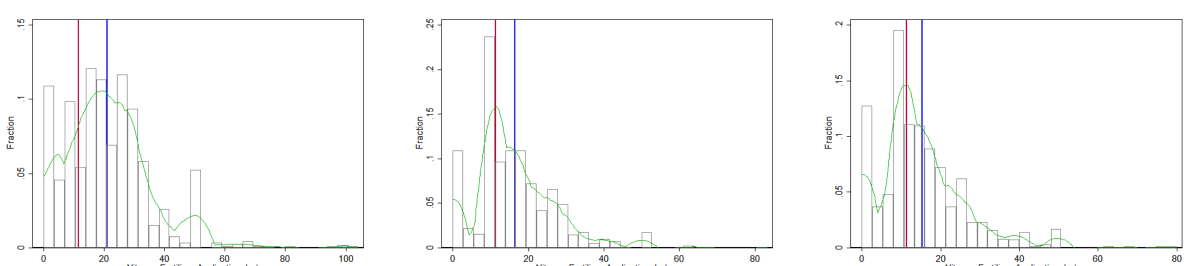
(a) Before Interventions, 2019 (b) After Phase 1 Intervention, 2020 (c) After Phase 2 Intervention, 2021



Note: This figure shows the distribution of nitrogen fertilizer application (kg/mu) in the planting stages. The red line indicates the mean of nitrogen recommendations, while the blue line shows the mean of actual use. The figure shows a clear pattern that there is no systematical difference in the planting stage.

Figure 12: Top-dressing Nitrogen Application in the Growing Stage [Used >> Recommended]

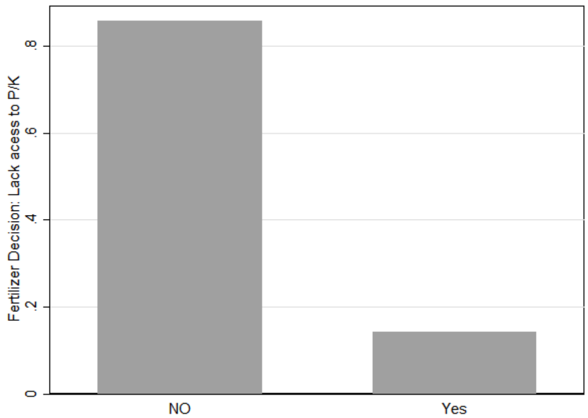
(a) Before Interventions, 2019 (b) After Phase 1 Intervention, 2020 (c) After Phase 2 Intervention, 2021



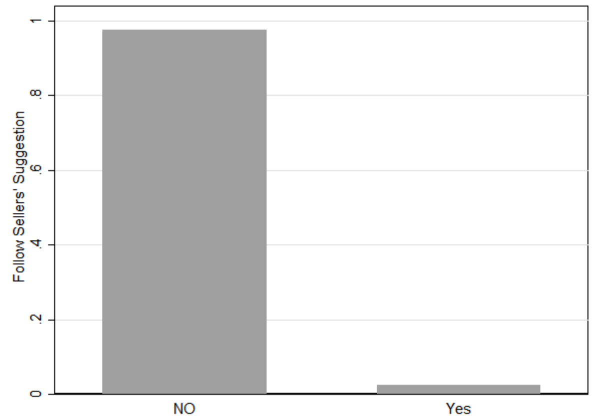
Note: This figure shows the distribution of nitrogen fertilizer application (kg/mu) in the growing stages. The red line indicates the mean of nitrogen recommendations, while the blue line shows the mean of actual use. The figure shows a clear pattern that nitrogen fertilizers are over-applied during the growing stages.

Figure 13: Supply Side-factors on Fertilizer Use

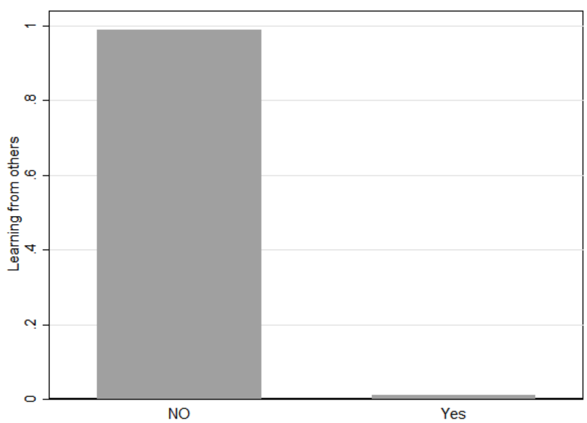
(a) Lack Access to P/K



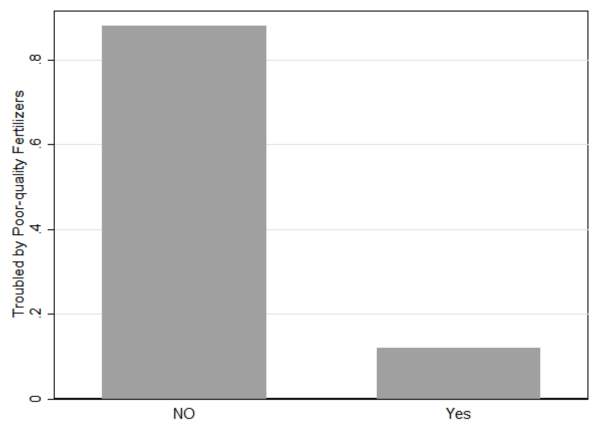
(b) Take Sellers' Suggestions



(c) Learning from Others (AEAs/Neighbors)



(d) Troubled by Poor-quality Fertilizers

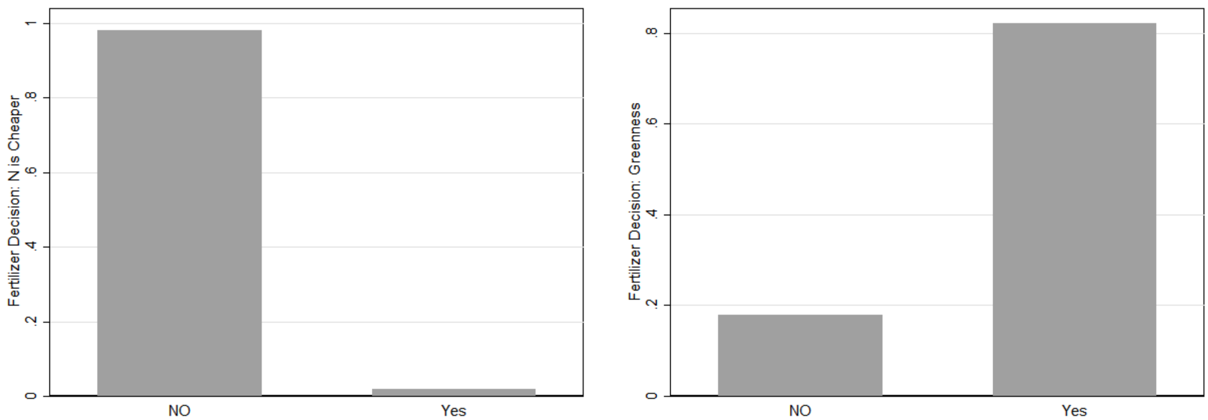


Note: The figure presents supply-side alternative explanations that might affect fertilizer usage. Figure 13a shows that farmers were not subject to the supply constraints of phosphorus and potassium. Figure 13b rejects the hypothesis that fertilizer sellers influenced farmers' choice of different fertilizer use. Only lower than 3% farmers followed the recommendations from sellers. Figure 13c suggests no evidence on learning from others, while Figure 13d shows direct evidence that quality was not a concern for farmers' choice.

Figure 14: Top-dressing Nitrogen Decision-making during the Growing Stage

(a) N Is Cheaper than P/K

(b) Greenness in Growing Stage



Note: The figure presents the fact that farmers' current amount of nitrogen application is not due to the lower price of nitrogen (Figure 14a). Figure 14b shows direct evidence on farmers' decision-making during the growing stage—based on greenness signals.

Table 1: Households' Characteristics, Fertilizer Recommendations, and Beliefs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$T1: ST$	$T2: App$	$T3: App + training$	$Control$	$T1 - C$	$T2 - C$	$T3 - C$
Panel A: Agricultural production							
Yield (kg/mu)	456.198	462.524	470.740	461.719	-5.522	0.805	9.020
	(108.445)	(109.091)	(104.811)	(107.706)	(8.824)	(8.851)	(8.677)
Area of plot (mu)	24.073	30.212	30.721	26.479	-2.405	3.733	4.242
	(69.053)	(88.619)	(91.842)	(74.921)	(5.883)	(6.700)	(6.843)
Revenue (RMB/mu)	1,095.754	1,095.747	1,121.603	1,120.591	-24.837	-24.844	1.012
	(391.229)	(382.665)	(428.841)	(421.841)	(33.217)	(32.883)	(34.730)
Profits (RMB/mu)	522.997	530.837	553.506	556.142	-33.145	-25.305	-2.636
	(473.205)	(472.774)	(545.048)	(493.563)	(39.477)	(39.460)	(42.453)
Compound fertilizer used (kg/mu)	37.140	36.032	35.636	35.748	1.392	0.284	-0.112
	(16.520)	(15.573)	(16.755)	(16.670)	(1.355)	(1.317)	(1.365)
Nitrogen fertilizer used (kg/mu)	19.676	20.774	22.256	21.177	-1.501	-0.404	1.079
	(14.000)	(14.700)	(15.158)	(13.936)	(1.140)	(1.169)	(1.189)
Phosphorus fertilizer used (kg/mu)	0.820	0.846	0.869	0.841	-0.021	0.005	0.027
	(5.561)	(6.098)	(6.771)	(5.779)	(0.463)	(0.485)	(0.514)
Share of HH using phosphorus fertilizer	0.027	0.027	0.023	0.023	0.003	0.003	-0.000
	(0.161)	(0.161)	(0.151)	(0.151)	(0.013)	(0.013)	(0.012)
Potassium fertilizer used (kg/mu)	1.235	1.230	1.263	1.236	-0.001	-0.005	0.027
	(4.945)	(6.593)	(5.996)	(4.966)	(0.405)	(0.477)	(0.450)
Share of HH using potassium fertilizer	0.090	0.090	0.090	0.087	0.003	0.003	0.003
	(0.287)	(0.287)	(0.287)	(0.282)	(0.023)	(0.023)	(0.023)
Total nitrogen used (kg/mu)	31.787	32.523	33.876	32.834	-1.047	-0.311	1.042
	(15.301)	(17.050)	(17.589)	(16.145)	(1.284)	(1.356)	(1.378)
Total phosphorus used (kg/mu)	15.105	14.704	14.575	14.590	0.515	0.114	-0.016
	(8.470)	(8.622)	(8.163)	(8.471)	(0.692)	(0.698)	(0.679)
Total potassium used (kg/mu)	13.615	13.241	13.141	13.151	0.463	0.090	-0.010
	(7.423)	(9.123)	(7.400)	(7.244)	(0.599)	(0.673)	(0.598)
Panel B: Recommendations based on soil testing							
Distance to the nearest testing plot (km)	0.276	0.301	0.266	0.315	-0.039	-0.014	-0.049*
	(0.305)	(0.283)	(0.314)	(0.351)	(0.027)	(0.026)	(0.027)
Compound recommendations	43.050	42.319	44.273	43.150	-0.101	-0.831	1.122*
(kg/mu)	(8.441)	(8.609)	(8.381)	(8.182)	(0.679)	(0.686)	(0.676)
Nitrogen fertilizer recommendations	11.440	11.572	11.215	11.398	0.042	0.174	-0.184
(kg/mu)	(3.691)	(3.668)	(2.938)	(3.049)	(0.276)	(0.275)	(0.244)
Phosphorus fertilizer recommendations	2.976	3.115	2.588	2.991	-0.016	0.124	-0.403
(kg/mu)	(4.892)	(4.997)	(4.299)	(5.990)	(0.447)	(0.450)	(0.426)
Potassium fertilizer recommendations	2.070	2.007	2.073	2.257	-0.187	-0.250	-0.184
(kg/mu)	(2.569)	(2.963)	(2.889)	(3.533)	(0.252)	(0.266)	(0.263)
Panel C: Beliefs							
Correct belief: greenness and yield	0.070	0.073	0.057	0.063	0.007	0.010	-0.007
	(0.256)	(0.261)	(0.232)	(0.244)	(0.020)	(0.021)	(0.019)
Correct belief: potassium and grain's density	0.020	0.017	0.017	0.017	0.003	0.000	-0.000
	(0.140)	(0.128)	(0.128)	(0.128)	(0.011)	(0.010)	(0.010)
Attrition rate	0.013	0.010	0.023	0.030	-0.017	-0.020*	-0.007
	(0.115)	(0.100)	(0.151)	(0.171)	(0.012)	(0.011)	(0.013)
Observations	300	300	300	300	600	600	600

Notes: Columns (5), (6), and (7) report the difference in characteristics between the treatment arms T1, T2, T3 and control groups, respectively. Standard deviations in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Effects of Different Treatment Arms on Fertilizer Usage by Timing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Aggregate Usage: (1)-(3)			**Planting Stage: (4)**	**Fertilizer Use in Growing Stage: (5)-(9)**				
Dept. Vars.	Total N	Total P	Total K	Compound Fertilizer	N-Fertilizer	P-Fertilizer	Share of HH Using P	K-Fertilizer	Share of HH Using K
T1: Soil Testing	-0.913	-0.259	-0.231	-0.705	-0.684	0.012	-0.004	0.004	-0.002
	(1.733)	(0.791)	(0.736)	(1.657)	(1.612)	(0.489)	(0.014)	(0.510)	(0.034)
T2: App	-3.924**	2.344***	1.374**	1.699	-4.478***	1.691***	0.239***	0.808*	0.241***
	(1.662)	(0.716)	(0.686)	(1.513)	(1.445)	(0.514)	(0.023)	(0.477)	(0.039)
T3: App +	-4.426***	2.721***	2.888***	2.457*	-5.227***	1.776**	0.232***	2.069***	0.324***
Training	(1.541)	(0.855)	(0.835)	(1.436)	(1.381)	(0.694)	(0.029)	(0.674)	(0.043)
Control Mean	30.84	14.74	13.31	36.16	19.05	0.833	0.0275	1.259	0.0893
Control SD	18.96	7.737	7.627	15.29	16.88	5.257	0.164	5.371	0.286
Clusters	200	200	200	200	200	200	200	200	200

Note: This table presents the treatment effect of the first-phase interventions on the application of different fertilizers in multiple stages. $T1$, $T2$, and $T3$ are indicators for three different treatment arms, indicating soil testing provision, soil testing + customized fertilizer recommendations through the app, and soil testing + customized fertilizer recommendations through the app + agricultural extension agents' training, respectively. Columns (1), (2), and (3) present treatment effects for the aggregate fertilizer application across all timings. Column (4) presents the treatment effect for N-P-K compound fertilizer use in the planting stage. Columns (5), (6) and (8) presents treatment effects for the application of top-dressing nitrogen-, phosphorus-, and potassium fertilizers in the growing stage. And columns (7) and (9) report the treatment effects for the proportion of households using phosphorus- and potassium fertilizers. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 3: Effects of First-phase Interventions on Yields, Profits, and Costs

	(1)	(2)	(3)	(4)	(5)	(6)
	First-phase Interventions					
Dept. Vars.	Yield	Log Yield	Profit	Revenue	Fertilizer	Other
VARIABLES	kg/mu		Yuan/mu	Yuan/mu	Cost	Cost
T1: Soil Testing	-7.523	-0.007	-1.244	-10.683	-6.286	-3.153
	(13.931)	(0.031)	(43.049)	(39.117)	(7.866)	(20.347)
T2: App	22.737*	0.054*	87.249**	68.569*	-7.048	-11.632
	(12.435)	(0.028)	(40.191)	(37.219)	(7.689)	(18.439)
T3: App + Soil Testing	31.647**	0.067**	82.420**	78.786**	3.452	-7.086
	(12.369)	(0.028)	(39.073)	(34.584)	(7.606)	(19.766)
Control Mean	465.6	6.117	520.7	1142	164.1	457.4
Control SD	106.9	0.262	360.8	298.2	83.48	168
Clusters	200	200	200	200	200	200

Note: This table presents the treatment effect of the first-phase interventions on the secondary outcomes. $T1$, $T2$, and $T3$ are indicators for three different treatment arms, indicating soil testing provision, soil testing + customized fertilizer recommendations through the app, and soil testing + customized fertilizer recommendations through the app + agricultural extension agents' training, respectively. Columns (1), (2), (3), and (4) present treatment effects for the yields, log yields, profits, and revenues. Columns (5) and (6) present the treatment effect for the costs of fertilizers and other inputs, respectively. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 4: Effects of First-phase Interventions on Beliefs about the Effectiveness

	(1)	(2)	(3)	(4)	(5)
	Beliefs about the effect of N/P/P/K: Correct = 1 (2)-(5)				
Dept. Vars.	Greenness & Yield	Nitrogen & Greenness	Phosphorus & Flowering Timing	Phosphorus & Root Length	Potassium & Grain's Density
T1: Soil Testing	0.006	-0.016	0.005	0.003	-0.004
	(0.021)	(0.021)	(0.035)	(0.019)	(0.016)
T2: App	0.150***	-0.033	0.220***	0.208***	0.171***
	(0.033)	(0.023)	(0.035)	(0.029)	(0.032)
T3: App + Training	0.310***	-0.024	0.340***	0.338***	0.368***
	(0.040)	(0.022)	(0.042)	(0.037)	(0.034)
Control Mean	0.0584	0.952	0.131	0.0412	0.0275
Control SD	0.235	0.214	0.338	0.199	0.164
Clusters	200	200	200	200	200

Note: This table presents the treatment effect of the first-phase interventions on farmers' beliefs. $T1$, $T2$, and $T3$ are indicators for three different treatment arms, indicating soil testing provision, soil testing + customized fertilizer recommendations through the app, and soil testing + customized fertilizer recommendations through the app + agricultural extension agents' training, respectively. Column (1) shows farmers' understanding on the relationship between greenness and yields. The outcome variable is a dummy for whether a farmer understood the relationship correctly (option 3, inverted-U shape relationship). Column (2) shows whether they understood the effects of nitrogen on greenness correctly. The outcome variables in columns (3), (4), and (5) are a set of dummies for whether farmers correctly understood the effects of phosphorus on flower timing, of phosphorus on root length, and the effects of potassium on grain's density, respectively. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 5: Second-phase Intervention: Balance between T1 and Control Group

	(1) *Control*	(2) *T1 : ST*	(3) *T1 − C*
Panel A: Agricultural production			
Yield (kg/mu)	465.554	458.031	-7.523
	(106.881)	(101.737)	(8.612)
Area of plot (mu)	27.290	26.263	-1.027
	(76.492)	(81.930)	(6.545)
Revenue (RMB/mu)	1,142.195	1,130.676	-11.519
	(298.177)	(285.161)	(24.079)
Profits (RMB/mu)	520.648	519.422	-1.225
	(360.765)	(352.065)	(29.422)
Compound fertilizer used (kg/mu)	36.157	35.451	-0.706
	(15.281)	(14.850)	(1.244)
Nitrogen fertilizer used (kg/mu)	19.051	18.368	-0.684
	(16.878)	(12.937)	(1.240)
Phosphorus fertilizer used (kg/mu)	0.833	0.845	0.012
	(5.257)	(5.727)	(0.454)
Proportion of HH using phosphorus fertilizer	0.027	0.024	-0.004
	(0.164)	(0.152)	(0.013)
Potassium fertilizer used (kg/mu)	1.260	1.262	0.002
	(5.373)	(4.834)	(0.422)
Proportion of HH using potassium fertilizer	0.089	0.088	-0.002
	(0.286)	(0.284)	(0.023)
Total nitrogen used (kg/mu)	30.842	29.928	-0.914
	(18.960)	(13.765)	(1.366)
Total phosphorus used (kg/mu)	14.739	14.480	-0.259
	(7.736)	(7.581)	(0.632)
Total potassium used (kg/mu)	13.313	13.079	-0.233
	(7.626)	(6.461)	(0.583)
Panel B: Recommendations based on soil testing			
Distance to the nearest testing plot (km)	0.312	0.278	-0.033
	(0.352)	(0.306)	(0.027)
Compound recommendations (kg/mu)	43.148	43.115	-0.033
	(8.119)	(8.354)	(0.680)
Nitrogen fertilizer recommendations (kg/mu)	11.381	11.404	0.023
	(2.954)	(3.621)	(0.273)
Phosphorus fertilizer recommendations (kg/mu)	3.009	2.960	-0.049
	(6.052)	(4.899)	(0.454)
Potassium fertilizer recommendations (kg/mu)	2.269	2.065	-0.204
	(3.557)	(2.567)	(0.256)
Panel C: Beliefs			
Correct belief: greenness and yield	0.058	0.064	0.006
	(0.235)	(0.246)	(0.020)
Correct belief: nitrogen and greenness	0.952	0.936	-0.016
	(0.214)	(0.246)	(0.019)
Correct belief: phosphorus and flowering timing	0.131	0.135	0.005
	(0.338)	(0.342)	(0.028)
Correct belief: phosphorus and root length	0.041	0.044	0.003
	(0.199)	(0.205)	(0.017)
Correct belief: potassium and grain's density	0.027	0.024	-0.004
	(0.164)	(0.152)	(0.013)
Attrition rate (second-follow survey)	0.014	0.027	0.013
	(0.117)	(0.162)	(0.012)
Observations	291	296	587

Notes: Standard deviations in parentheses for all the columns.Column (3) reports the difference in characteristics between the treatment arm T1 and control groups using the data from the first follow-up survey in November 2020.

Table 6: Effects of Leaf Color Charts on Fertilizer Usage in Different Timings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Aggregate Usage: (1)-(3)			**Planting Stage: (4)**	**Fertilizer Use in Growing Stage: (5)-(9)**				
Dept. Vars.	Total N	Total P	Total K	Compound Fertilizer	N-Fertilizer	P-Fertilizer	Share of HH Using P	K-Fertilizer	Share of HH Using K
Panel A: Phase 2 Interventions for comparison									
T1: LCC	-3.757**	0.482	0.173	0.223	-3.829**	0.396	0.062***	0.098	0.066*
	(1.662)	(0.828)	(0.745)	(1.750)	(1.519)	(0.521)	(0.021)	(0.560)	(0.040)
T2: App + LCC	-4.865***	2.187***	1.898**	1.891	-5.482***	1.459***	0.206***	1.267*	0.247***
	(1.631)	(0.815)	(0.865)	(1.564)	(1.410)	(0.523)	(0.025)	(0.690)	(0.038)
T3: App +	-4.809***	2.917***	3.143***	2.951**	-5.771***	1.782**	0.246***	2.159***	0.316***
Training + LCC	(1.620)	(0.914)	(0.851)	(1.465)	(1.466)	(0.737)	(0.033)	(0.687)	(0.045)
Control Mean	30.48	14.42	13.18	35.07	19.04	0.928	0.0418	1.488	0.0941
Control SD	19.27	8.347	8.337	15.72	17.12	5.799	0.201	6.080	0.292
Clusters	199	199	199	199	199	199	199	199	199
Panel B: Recap of Phase 1 Interventions									
T1: Soil Testing	-0.913	-0.259	-0.231	-0.705	-0.684	0.012	-0.004	0.004	-0.002
	(1.733)	(0.791)	(0.736)	(1.657)	(1.612)	(0.489)	(0.014)	(0.510)	(0.034)
T2: App	-3.924**	2.344***	1.374**	1.699	-4.478***	1.691***	0.239***	0.808*	0.241***
	(1.662)	(0.716)	(0.686)	(1.513)	(1.445)	(0.514)	(0.023)	(0.477)	(0.039)
T3: App + Training	-4.426***	2.721***	2.888***	2.457*	-5.227***	1.776**	0.232***	2.069***	0.324***
	(1.541)	(0.855)	(0.835)	(1.436)	(1.381)	(0.694)	(0.029)	(0.674)	(0.043)

Note: Panel A presents the treatment effect of the second-phase intervention intertwined with the first-phase interventions on the application of different fertilizers in multiple stages. $T1$, $T2$, and $T3$ are indicators for three different treatment arms, indicating leaf color charts, customized fertilizer recommendations through the app + leaf color charts, and customized fertilizer recommendations through the app + agricultural extension agents' training + leaf color charts, respectively. Columns (1), (2), and (3) present treatment effects for the aggregate fertilizer application across all timings. Column (4) presents the treatment effect for N-P-K compound fertilizer use in the planting stage. Columns (5), (6) and (8) presents treatment effects for the application of top-dressing nitrogen-, phosphorus-, and potassium fertilizers in the growing stage. And columns (7) and (9) report the treatment effects for the proportion of households using phosphorus- and potassium fertilizers.
For panel B, we replicate the treatment effect of the first-phase interventions on the application of different fertilizers in multiple stages. $T1$, $T2$, and $T3$ are indicators for three different treatment arms, indicating soil testing provision, soil testing + customized fertilizer recommendations through the app, and soil testing + customized fertilizer recommendations through the app + agricultural extension agents' training, respectively. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 7: Effects of Second-phase Interventions on Beliefs about the Effectiveness

	(1)	(2)	(3)	(4)	(5)
		Beliefs about the effect of N/P/P/K: Correct = 1 (2)-(5)			
Dept. Vars.	Greenness & Yield	Nitrogen & Greenness	Phosphorus & Flowering Timing	Phosphorus & Root Length	Potassium & Grain's Density
T1: LCC	0.271***	-0.010	0.038	0.031	0.031
	(0.038)	(0.021)	(0.033)	(0.020)	(0.024)
T2: App + LCC	0.375***	-0.031	0.247***	0.224***	0.203***
	(0.040)	(0.023)	(0.034)	(0.030)	(0.035)
T3: App + Training + LCC	0.444***	-0.028	0.350***	0.340***	0.354***
	(0.040)	(0.023)	(0.041)	(0.037)	(0.034)
Control Mean	0.0627	0.951	0.115	0.0383	0.0383
Control SD	0.243	0.216	0.320	0.192	0.192
Clusters	199	199	199	199	199

Note: This table presents the treatment effect of the first-phase interventions on farmers' beliefs. $T1$, $T2$, and $T3$ are indicators for three different treatment arms, indicating leaf color charts, customized fertilizer recommendations through the app + leaf color charts, and customized fertilizer recommendations through the app + agricultural extension agents' training + leaf color charts, respectively. Column (1) shows farmers' understanding on the relationship between greenness and yields. The outcome variable is a dummy for whether a farmer understood the relationship correctly (option 3, inverted-U shape relationship). Column (2) shows whether they understood the effects of nitrogen on greenness correctly. The outcome variables in columns (3), (4), and (5) are a set of dummies for whether farmers correctly understood the effects of phosphorus on flower timing, of phosphorus on root length, and the effects of potassium on grain's density, respectively. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 8: Effects of Two-phase Interventions on Gap between Applications and Recommendations

	(1)	(2)	(3)	(4)	(5)	(6)
	First-phase Intervention (1) - (3)			***Second-phase Intervention (4) - (6)***		
Dept. Vars.	Nitrogen Gap	Phosphorus Gap	Potassium Gap	Nitrogen Gap	Phosphorus Gap	Potassium Gap
	[Gap = the Used - the Recommended]					
T1: Soil Testing/LCC	-0.740	-0.197	-0.019	-3.581**	0.560	0.388
	(1.797)	(0.990)	(0.784)	(1.724)	(1.018)	(0.788)
T2: App/LCC + App	-3.614**	2.212**	1.815**	-4.500***	2.375**	2.401***
	(1.690)	(0.965)	(0.742)	(1.648)	(0.997)	(0.859)
T3: App + Training/App + Training + LCC	-4.661***	2.936***	2.688***	-5.109***	3.070***	2.913***
	(1.615)	(1.018)	(0.913)	(1.710)	(1.098)	(0.918)
Observations	1177	1177	1177	1153	1153	1153
Control Mean	7.637	-4.866	-3.339	7.269	-5.205	-3.465
Control SD	19.34	10.13	8.559	19.63	10.45	9.160
Clusters	200	200	200	199	199	199
R squared	0.0167	0.0180	0.0186	0.0175	0.0141	0.0182

Note: Columns (1), (2) and (3) present the treatment effect of the first-phase interventions on the gap in fertilizers between the actual application and recommended use. $T1$, $T2$, and $T3$ are indicators for three different treatment arms, indicating soil testing provision, soil testing + customized fertilizer recommendations through the app, and soil testing + customized fertilizer recommendations through the app + agricultural extension agents' training, respectively. The outcome variables in column (1), (2), and (3) are nitrogen use gap [used - recommended], phosphorus use gap [used - recommended], and potassium use gap [used - recommended], respectively.

Columns (4), (5) and (6) present the treatment effect of the second-phase interventions on the gap in fertilizers between the actual application and recommended use. $T1$, $T2$, and $T3$ are indicators for three different treatment arms, indicating leaf color charts, customized fertilizer recommendations through the app + leaf color charts, and customized fertilizer recommendations through the app + agricultural extension agents' training + leaf color charts, respectively. The outcome variables in columns (4), (5), and (6) are nitrogen use gap [used - recommended], phosphorus use gap [used - recommended], and potassium use gap [used - recommended] in the second follow-up survey, respectively. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 9: IV Estimation: Deviation in Fertilizer Application and Yields

	(1)	(2)	(3)
	IV-First Stage	2SLS	2SLS
	logGap^2	Yields	Log Yields
T2 (App)	-1.029***		
	(0.191)		
T3 (App + AEA's Training)	-1.134***		
	(0.223)		
LogGap^2		-36.39***	-0.0775**
		(10.86)	(0.0246)
Observations	465	465	462
R-squared	0.124		
Control Mean	5.27	466.58	6.13
F-statistic	23.48		

Note: In this table, we employ the IV-2sls regression strategy using data from the second round of survey to estimate the impact of deviation in fertilizer application compared to the recommended use on yields. In the IV-2sls regression, we use $T2$ and $T3$ indicators as the instrumental variables to run the equations (5) and (6). We also limit the regression samples to those who overuse nitrogen fertilizers and underuse phosphorus/potassium fertilizers so that the underlying relationships between fertilizer gaps and yields are clearly defined. The overuse of nitrogen and underuse of phosphorus/potassium directly lead to lower yield. Column (1) reports the first-stage regression and the outcome variable is the log of Gap^2, defined as $Gap^2 = (N\_Gap)^2 + (P\_Gap)^2 + (K\_Gap)^2$. The outcome variables in columns (2) and (3) are yields and the log of yields. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 10: IV Estimation: Deviation in Fertilizer Application and Revenues/Costs

	(1)	(2)	(3)	(4)
	IV-First Stage	2SLS	2SLS	2SLS
	logGap^2	Revenues	Fertilizer Cost	Other Cost
T2 (App)	-1.029***			
	(0.191)			
T3 (App + AEA's Training)	-1.134***			
	(0.223)			
LogGap^2		-98.20**	3.251	-24.16
		(33.81)	(4.992)	(17.26)
Observations	465	465	465	465
R-squared	0.124			
Control Mean	5.27	1147.15	162.84	480.76
F-statistic	23.48			

Note: In this table, we employ the IV-2sls regression strategy using data from the second round of survey to estimate the impact of deviation in fertilizer application compared to the recommended use on revenues, fertilizer costs, and other costs. In the IV-2sls regression, we use $T2$ and $T3$ indicators as the instrumental variables to run the equations (5) and (6). We also limit the regression samples to those who overuse nitrogen fertilizers and underuse phosphorus/potassium fertilizers so that the underlying relationships between fertilizer gaps and these outcome variables are clearly defined. The overuse of nitrogen and underuse of phosphorus/potassium directly lead to lower revenues. Column (1) reports the first-stage regression and the outcome variable is the log of Gap^2, defined as $Gap^2 = (N\_Gap)^2 + (P\_Gap)^2 + (K\_Gap)^2$. The outcome variables in columns (2), (3) and (4) are revenues, fertilizer costs, and other costs. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Appendix A: Supplementary Tables and Figures

Figure A1: Nitrogen Abuse in Developed Countries

With Too Much of a Good Thing, Europe Tackles Excess Nitrogen

In Germany, the Netherlands, Denmark and other countries, European governments are beginning to push farmers, industry, and municipalities to cut back on fertilizers and other sources of nitrogen that are causing serious environmental harm.

BY CHRISTIAN SCHWÄGERL · APRIL 14, 2015

Only seconds after Claudia Wiedner drops the metallic rod into the gray waters of Lake Scharmützel, 30 miles southeast of Berlin, the probe starts sending signals back to her computer. On a cold, foggy day in March, Wiedner, a limnologist at the Brandenburg University of Cottbus-Senftenberg, and a research technician are out on the water in their small vessel to investigate nitrogen pollution.

The New York Times

Polluting Farmers Should Pay

Nitrogen and phosphorus pollution, commonly called nutrient pollution, the bulk of which comes from agricultural fertilizer and manure runoff. ... This may sound like a lot, but five times that was spent on industrial and

Aug 25, 2019

The New York Times

Killer Slime, Dead Birds, an Expunged Map: The Dirty Secrets ...

The map juxtaposed pollution in northern Italy with the European Union ... The New York Times created an approximation that confirms what ...

Dec 25, 2019

The New York Times

Fertilizers, a Boon to Agriculture, Pose Growing Threat to U.S. Waterways

Nitrogen-based fertilizers, which came into wide use after World War II, ... this form of pollution, leading to more damaging algae blooms and dead zones in American coastal waters. ... Michael Kirby Smith for The New York

Jul 27, 2017

Australia's Nutrient Pollution Travels from River to Reef

Meanwhile, in the southern hemisphere, nutrient pollution from nearly 40 river basins exacerbates climate change's threat to the Great Barrier Reef. Nutrients in the coastal waters impair the reef's resilience during bleaching events and trigger harmful algal blooms that feed the reef-eating crown-of-thorns starfish.

According to the government, the country needs to reduce nutrient pollution by 80 percent, primarily from farms. The government passed laws restricting land-use changes in the hopes of reducing runoff, but without involvement from agricultural stakeholders, buy-in around the greater nutrient-reduction effort has been limited.

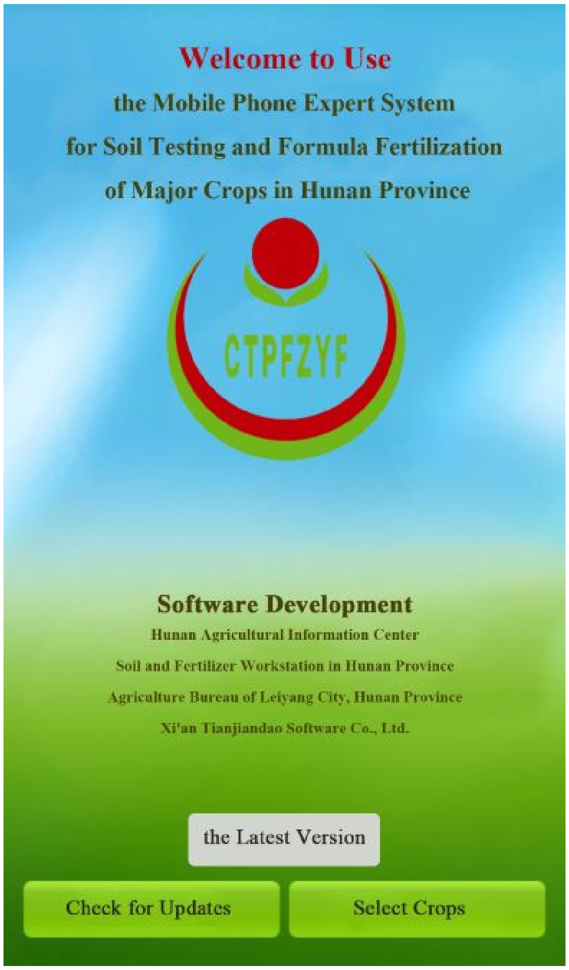
Stepping-up Global Action to Address the Nutrient Challenge

On March 11-15, the highest-level decision-making body on the environment will convene in Nairobi for the fourth session of the UN Environment Assembly (UNEA4). Leaders and high-level decisionmakers representing the UN's 192 member states will discuss intergovernmental cooperation around environmental goals and policies, including for water pollution. We expect a first-of-its-kind resolution on

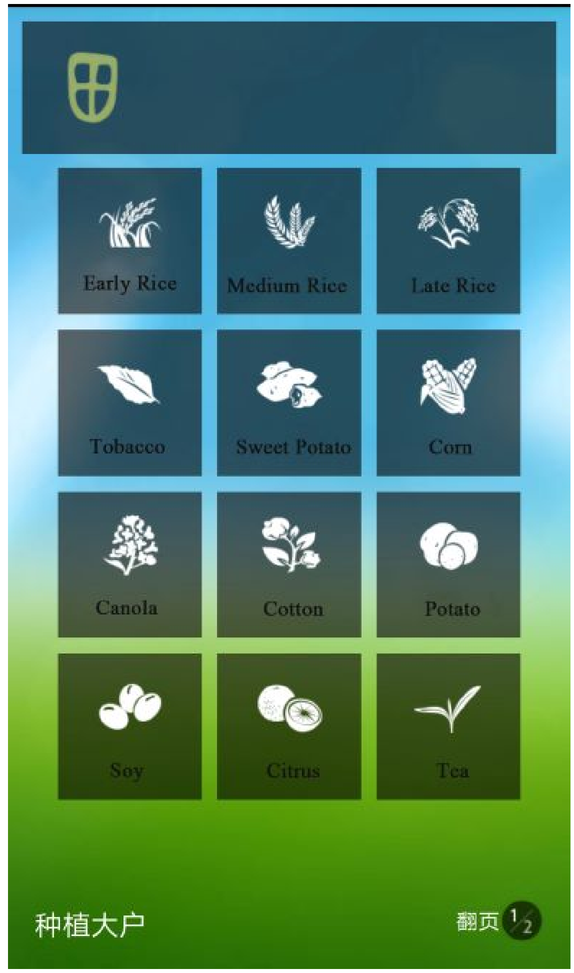
Note: Agricultural nitrogen fertilizer use has become one of the major sources for N_2O pollution.

Figure A2: The First Two Interfaces of the Mobile Application

(a) Endorsed by Hunan Government



(b) Scalable to Up to 15 Crops



Note: The left panel shows that the mobile application is endorsed by Hunan government. The right panel asks farmers to choose crops to get fertilizer recommendations.

Figure A3: The Second Two Pages of the Mobile Application

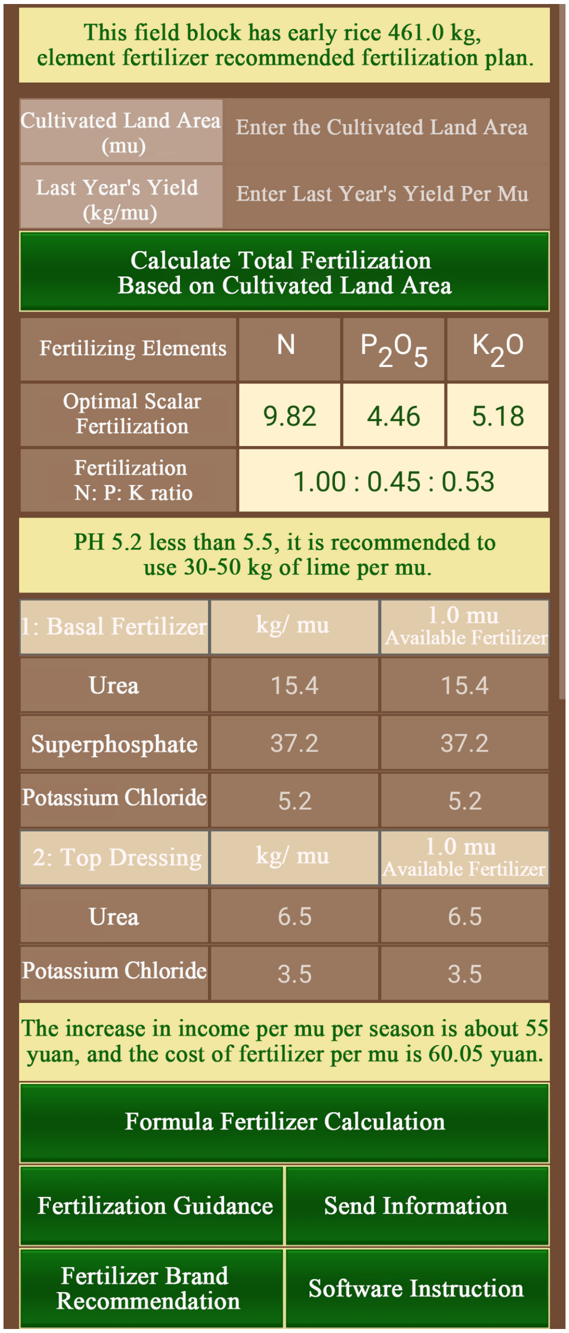
(a) Acquire Soil Analysis by GPS Tracking

(b) Acquire Soil Analysis by Selecting Places

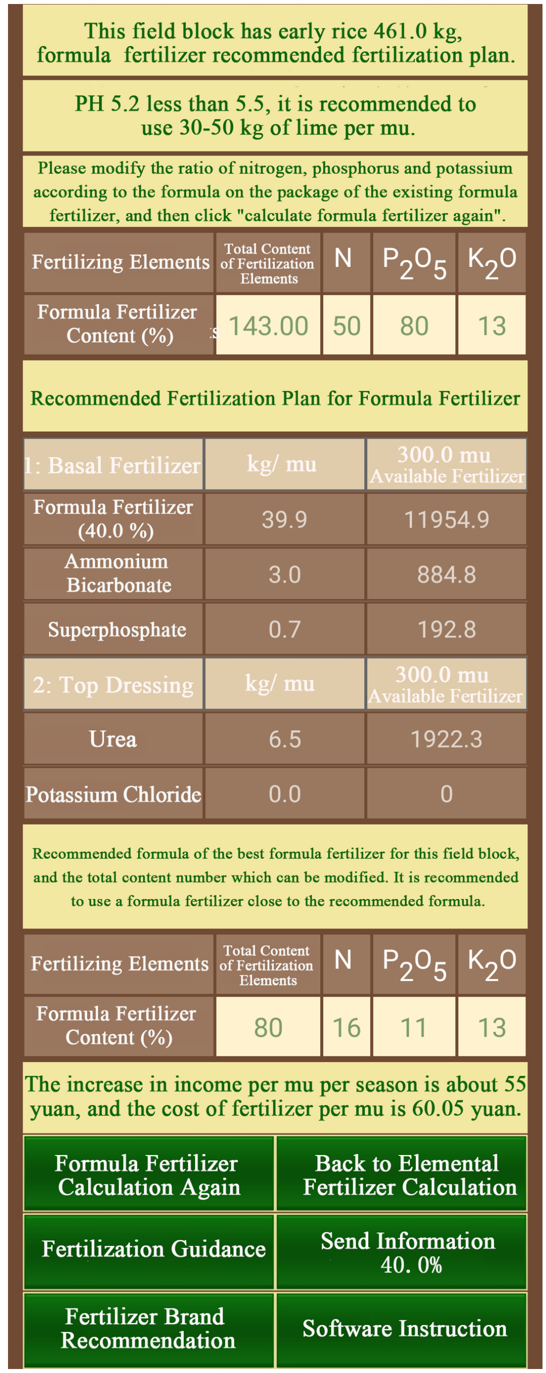
Note: The left panel shows that farmers can acquire soil testing data by GPS tracking or by choosing locations. The app then displays the amount of pH value, organic matter, nitrogen, phosphorus, and potassium in farmers' plots. The right panel shows a set of locations that farmers can select from.

Figure A4: Dynamic Fertilizer Recommendations

(a) Combination of Different Individual Fertilizers



(b) Combination of Compound + Other Individual Fertilizers

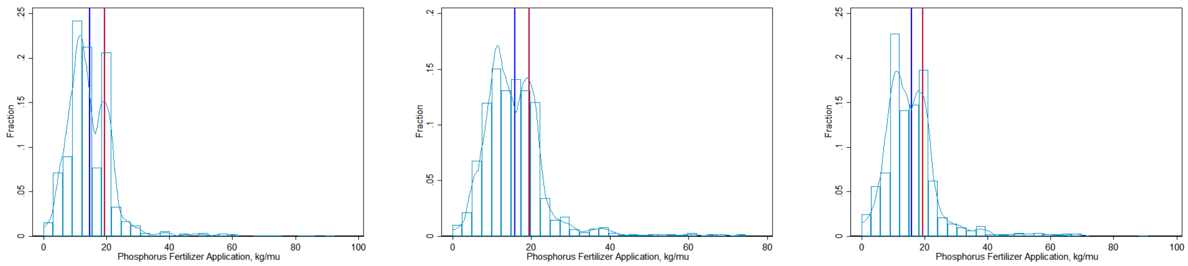


Note: The left panel shows that the app can display the customized recommendations of different individual fertilizers for different timing based on personalized soil testing. Since most farmers are using the compound fertilizers, the right panel shows that the app can display the customized recommendations of the combination of compound fertilizers and individual fertilizers for different timing based on personalized soil testing.

Figure A5: Total Phosphorus Application [Used < Recommended]

(Red = Mean Recommendations, blue = Mean Application)

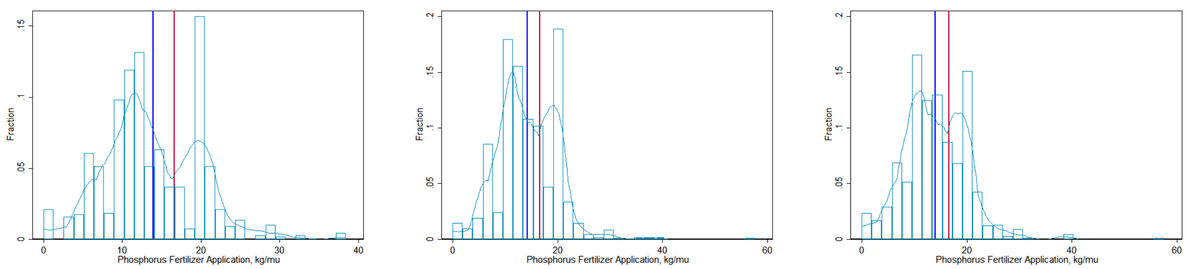
Before Interventions 2019 | After Phase 1 Intervention, 2020 | After Phase 2 Intervention, 2021



Note: This figure shows the distribution of total phosphorus fertilizer application (kg/mu) throughout the cropping cycle. The red line indicates the mean of phosphorus recommendations, while the blue line shows the mean of actual use. The figure suggests a clear pattern of phosphorus underuse and our interventions reduced such gap.

Figure A6: Phosphorus Application in the Planting Stage [Used < Recommended]

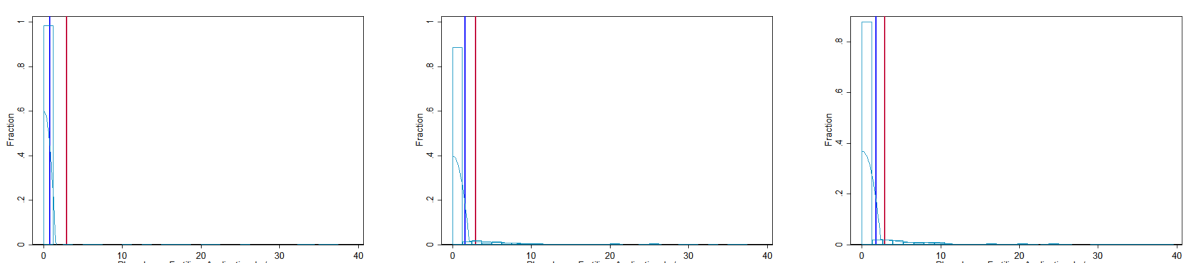
(a) Before Interventions, 2019 (b) After Phase 1 Intervention, 2020 (c) After Phase 2 Intervention, 2021



Note: This figure shows the distribution of top-dressing phosphorus fertilizer application (kg/mu) in the planting stages. The red line indicates the mean of phosphorus recommendations, while the blue line shows the mean of actual use. The figure suggests a clear pattern of phosphorus underuse during the planting stages.

Figure A7: Top-dressing Phosphorus Application in the Growing Stage [Used < Recommended]

(a) Before Interventions, 2019 (b) After Phase 1 Intervention, 2020 (c) After Phase 2 Intervention, 2021

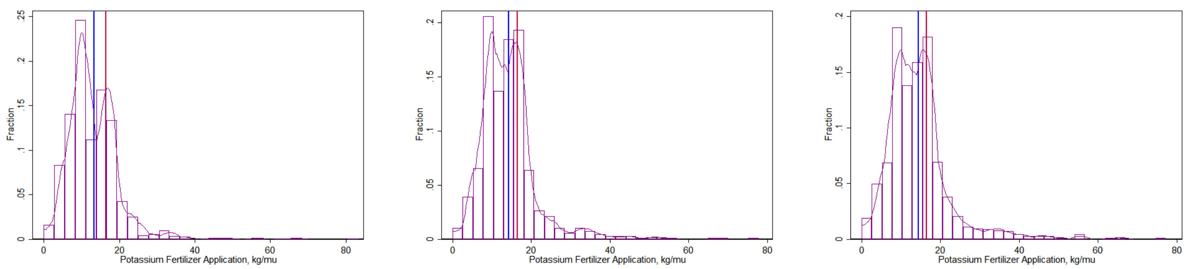


Note: This figure shows the distribution of phosphorus fertilizer application (kg/mu) in the growing stages. The red line indicates the mean of phosphorus recommendations, while the blue line shows the mean of actual use. The figure suggests underuse of phosphorus in the growing stages, and our interventions reduced such gap.

Figure A8: Total Potassium Application [Used < Recommended]

(Red = Mean Recommendations, blue = Mean Application)

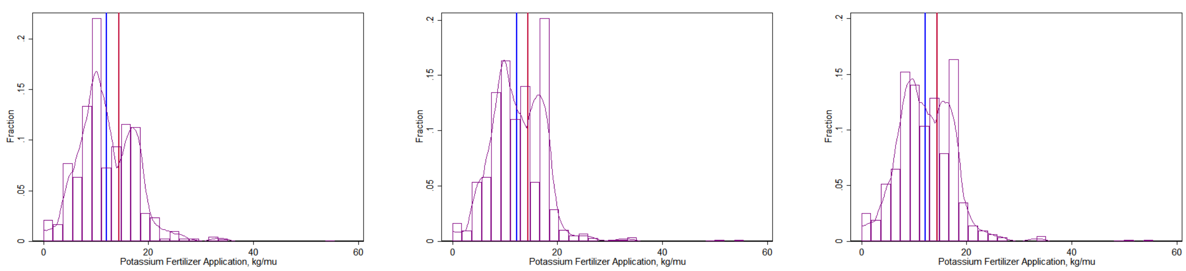
Before Interventions 2019 | After Phase 1 Intervention, 2020 | After Phase 2 Intervention, 2021



Note: This figure shows the distribution of total potassium fertilizer application (kg/mu) throughout the cropping cycle. The red line indicates the mean of potassium recommendations, while the blue line shows the mean of actual use. The figure suggests a clear pattern of potassium underuse and our interventions reduced such gap.

Figure A9: Potassium Application in the Planting Stage [Used < Recommended]

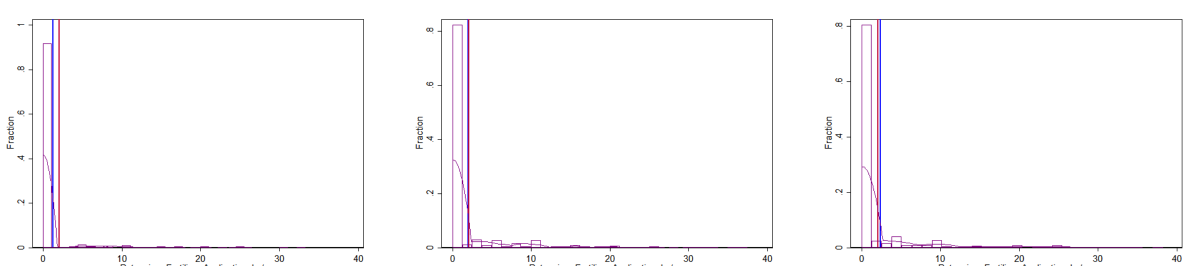
(a) Before Interventions 2019 | (b) After Phase 1 Intervention, 2020 | (c) After Phase 2 Intervention, 2021



Note: This figure shows the distribution of potassium fertilizer application (kg/mu) in the planting stages. The red line indicates the mean of potassium recommendations, while the blue line shows the mean of actual use. The figure suggests underuse of potassium in the planting stages.

Figure A10: Top-dressing Potassium Application in the Growing Stage [Used < Recommended]

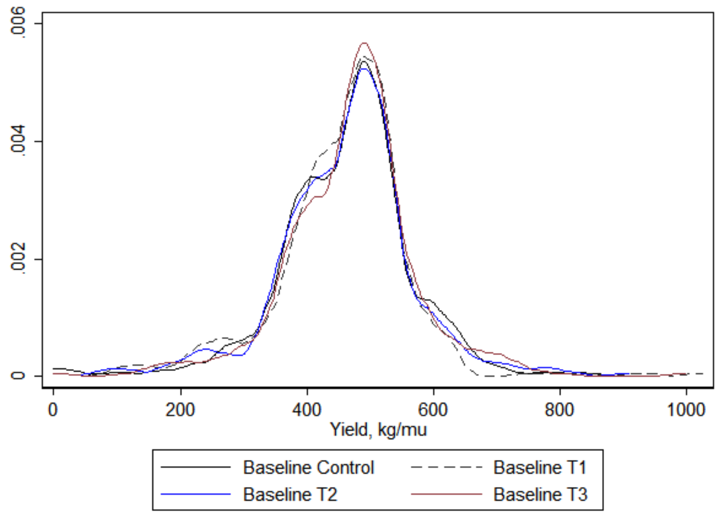
(a) Before Interventions 2019 | (b) After Phase 1 Intervention, 2020 | (c) After Phase 2 Intervention, 2021



Note: This figure shows the distribution of potassium fertilizer application (kg/mu) in the growing stages. The red line indicates the mean of potassium recommendations, while the blue line shows the mean of actual use. The figure suggests underuse of potassium in the growing stages, and our interventions reduced such gap.

Figure A11: Yields before/after the First-phase Interventions

(a) Yields at Baseline (kg/mu)



(b) Yields after First-phase Interventions (kg/mu)

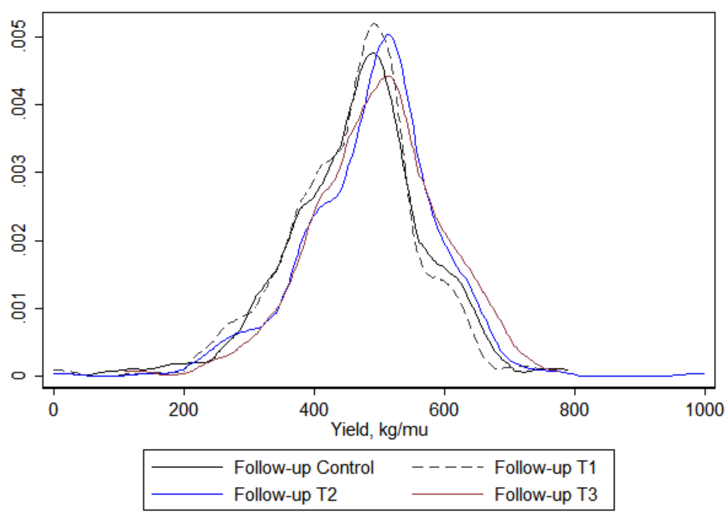


Table A1: Distance and Fertilizer Gap between Applications and Recommendations

	(1)	(2)	(3)	(4)	(5)	6
	Baseline (1) - (3)			***First Follow-up (4) - (6)***		
Dept. Vars.	Nitrogen Gap	Phosphorus Gap	Potassium Gap	Nitrogen Gap	Phosphorus Gap	Potassium Gap
	[Gap = the Used - the Recommended]					
Distance to the Nearest Testing Point	0.992	0.088	0.297	1.951	-0.103	0.222
	(1.335)	(1.198)	(0.819)	(1.307)	(0.961)	(0.882)
Observations	1200	1200	1200	1177	1177	1177
Control Mean	9.632	-4.997	-3.489	7.637	-4.866	-3.339
Control SD	16.46	9.762	8.192	19.34	10.13	8.559
Clusters	200	200	200	200	200	200
R squared	0.000344	7.46e-06	0.000122	0.00163	9.77e-06	6.42e-05

Note: Columns (1), (2) and (3) present the relationship between the distance to the nearest soil testing plots and the gap in fertilizers between the actual application and recommended use using the baseline data. The outcome variables in column (1), (2), and (3) are nitrogen use gap [used - recommended], phosphorus use gap [used - recommended], and potassium use gap [used - recommended], respectively.

Columns (4), (5) and (6) present the relationship between the distance to the nearest soil testing plots and the gap in fertilizers between the actual application and recommended use using the first follow-up data. The outcome variables in columns (4), (5), and (6) are nitrogen use gap [used - recommended], phosphorus use gap [used - recommended], and potassium use gap [used - recommended] in the second follow-up survey, respectively. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table A2: Validity of IVs: using T2 and T3 Indicators before the Interventions

	(1)	(2)	(3)
	IV-First Stage	2SLS	2SLS
	(1)	(2)	(3)
	logGap^2	Yields	Log Yields
T2 (App)	-0.0716		
	(0.104)		
T3 (App + AEA's Training)	-0.00594		
	(0.106)		
LogGap^2		-79.60	-0.0655
		(189.1)	(0.402)
Observations	568	568	564
R-squared	0.000853	.	.
Control Mean	5.66	460.03	6.11
F-statistic	0.252		

Note: In this table, we replicate the IV-2sls regression using $T2$ and $T3$ indicators in the baseline data and employing equations (5) and (6). We also limit the regression samples to those who overuse nitrogen fertilizers and underuse phosphorus/potassium fertilizers so that the underlying relationships between fertilizer gaps and yields are clearly defined. The overuse of nitrogen and underuse of phosphorus/potassium directly lead to lower yield. Column (1) reports the first-stage regression and the outcome variable is the log of Gap^2, defined as $Gap^2 = (N\_Gap)^2 + (P\_Gap)^2 + (K\_Gap)^2$. The outcome variables in columns (2) and (3) are yields and the log of yields. We do not find any significance either in the first-stage or second-stage regressions since $T2$ and $T3$ did not affect the fertilizer applications and yields in the baseline data. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table A3: IV Estimation II: Using T2 and T3 Indicators as IVs for Each Fertilizer Gap

	(1)	(2)	(3)	(4)	(5)	(6)
Dept. Vars.	Yields	Yields	Yields	Yields	Yields	Yields
Nitrogen Gap	-8.243***					
	(2.942)					
Phosphorus Gap		11.615***				
		(4.237)				
Potassium Gap			13.466***			
			(5.007)			
Nitrogen Gap Ratio				-1.906***		
				(0.726)		
Phosphorus Gap Ratio					2.362**	
					(0.947)	
Potassium Gap Ratio						2.108***
						(0.776)
Observations	1,177	1,177	1,177	1,177	1,177	1,177

Note: In this table, we conduct another type of IV regression. In columns (1), (2) and (3), we instrument the gap in nitrogen use [Used - Recommended], gap in phosphorus use, and gap in potassium use separately with the $T2$ and $T3$ indicators. We then present the second-stage estimation of the effects of gap on yields separately in these columns.

In columns (4), (5) and (6), we instrument the gap ratio in nitrogen use [(Used - Recommended)/Recommended], gap ratio in phosphorus use, and gap ratio in potassium use separately with the $T2$ and $T3$ indicators. We then present the second-stage estimation of the effects of gap on yields separately in these columns. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table A4: IV Estimation: Deviation in Fertilizer Application and Yields

	(1)	(2)	(3)
	IV-First Stage	2SLS	2SLS
	$logGap^2$	Yields	Log Yields
T2 (App)	-0.664***		
	(0.149)		
T3 (App + AEA's Training)	-0.739***		
	(0.155)		
$LogGap^2$		-44.57**	-0.0934**
		(13.87)	(0.0309)
Observations	1177	1177	1173
R-squared	0.0498		
Control Mean	5.41	465.55	6.12
F-statistic	17.88		

Note: In this table, we employ the IV-2sls regression strategy using data from the second round of survey to estimate the impact of deviation in fertilizer application compared to the recommended use on yields. Different from Table 9, we use all the observations from the second follow-up survey here. In the IV-2sls regression, we use $T2$ and $T3$ indicators as the instrumental variables to run the equations (5) and (6). Column (1) reports the first-stage regression and the outcome variable is the log of Gap^2, defined as $Gap^2 = (N\_Gap)^2 + (P\_Gap)^2 + (K\_Gap)^2$. The outcome variables in columns (2) and (3) are yields and the log of yields. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table A5: IV Estimation: Deviation in Fertilizer Application and Revenues/Costs

	(1)	(2)	(3)	(4)
	IV-First Stage	2SLS	2SLS	2SLS
	$logGap^2$	Revenues	Fertilizer Cost	Other Cost
T2 (App)	-0.664***			
	(0.149)			
T3 (App + AEA's Training)	-0.739***			
	(0.155)			
$LogGap^2$		-112.8**	-1.562	10.69
		(38.15)	(7.621)	(20.83)
Observations	1177	1177	1177	1177
R-squared	0.0498			
Control Mean	5.41	1142.20	164.17	457.37
F-statistic	17.88			

Note: In this table, we employ the IV-2sls regression strategy using data from the second round of survey to estimate the impact of deviation in fertilizer application compared to the recommended use on revenues, fertilizer costs, and other costs. Different from Table 10, we use all the observations from the second follow-up survey here. In the IV-2sls regression, we use $T2$ and $T3$ indicators as the instrumental variables to run the equations (5) and (6). The overuse of nitrogen and underuse of phosphorus/potassium directly lead to lower revenues. Column (1) reports the first-stage regression and the outcome variable is the log of Gap^2, defined as $Gap^2 = (N\_Gap)^2 + (P\_Gap)^2 + (K\_Gap)^2$. The outcome variables in columns (2), (3) and (4) are revenues, fertilizer costs, and other costs. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Appendix B: Proofs.

In this section we detail the derivation described in Section 4.3. Let's begin with equation (4).

After observing profit π_t contributed by the realized states b_{Nt} and b_{Kt}, the farmer believes that the realized state was b_{Nt} and $\tilde{b}_{Kt}$ satisfying:

$$\tilde{\Pi}\left(a_{Nt}, a_{Kt}, b_{Nt}, \tilde{b}_{Kt}\right) = \pi_t = \Pi\left(a_{Nt}, a_{Kt}, b_{Nt}, b_{Kt}\right)$$

At $\Pi = \tilde{\Pi}$, $\tilde{b_K}$ serves as the intermediate to connect the following two terms:

$$(1-\lambda) f_1(a_N) \exp(b_N) + f_2(a_K) \exp(b_K) = f_2(a_K) \exp\left(\tilde{b}_K\right)$$

$$\tilde{b}_K(a_N, a_K, b_K) = \tilde{b}_K(a_K, b_K) = \log\left(C + f_2(a_K) \exp(b_K)\right) - \log\left(f_2(a_K)\right) \tag{7}$$

Where $C = (1-\lambda) f_1(a_N) \exp(b_N)$. Hence we could derive the distorted belief $\tilde{b}_K$.

Turning back to the true model, the optimal action a_K^* satisfies: $f_2'(a_K^*) \exp\left(\theta_K + \sigma_K^2/2\right) = c_K$. By taking logs on both sides and rearranging terms, we obtain $\log\left(f_2'(a_K^*)\right) = \log(c_K) - \theta_K - \sigma_K^2/2$

As for the misspecified mode:

$$\log\left(f_2'(\tilde{a}_K^*)\right) = \log(c_K) - \tilde{\theta}_K - \sigma_K^2/2 \tag{8}$$

Again for $\tilde{b}_K(a_K, b_K) = \log\left(C + f_2(a_K) \exp(b_K)\right) - \log\left(f_2(a_K)\right)$, because $\lambda > 1$, $f_1(a_N) > 0$, then $C < 0$. Thus then we can always find a negative number A to obtain

$$\log\left(C + f_2(a_K) \exp(b_K)\right) = A + \log\left(f_2(a_K) \exp(b_K)\right)$$

Convergence of farmer's beliefs. Heidhues, Kőszegi, and Strack (2021) define the function $g: \mathbb{N} \times \mathbb{R} \rightarrow \mathbb{R}$ as the objective expectation of $\tilde{b}_{t+1} - \tilde{\theta}_t$:

$$g(t, x) = \mathbb{E}\left[\tilde{b}\left(b_{t+1}, a^*(t+1, x)\right)\right] - x$$

Then how are farmer's subjective beliefs updated in the limit is defined as:

$$g(x) = \lim_{t \rightarrow \infty} g(t, x) = \mathbb{E}\left[\tilde{b}\left(b_{t+1}, a^*(x)\right)\right] - x$$

According to such expectation of the updating convergence function, we obtain:

$$g(x) = \lim_{t \rightarrow \infty} E\left(\tilde{b}_K\left(b_K, a_K^*(x)\right)\right) - \tilde{\theta}_K \tag{9}$$

Where $\tilde{\theta}_K$ is the mean of incorrectly specified belief regarding the effectiveness of phosphorus (P)/potassium (K). As $t \rightarrow \infty$, $g(a_K^*) = A + \log\left(f_2(a_K^*)\right) + E(b_K) - \log\left(f_2(a_K^*)\right) - \tilde{\theta}_K$, and from

Equation (8) above, we also have that

$$\tilde{\theta}_K = \log\left(c_K\right) - \log\left(f_2'\left(\tilde{a}_K^*\right)\right) - \sigma_K^2/2$$

Substitute it into the g function in Equation (9), and solve $g = 0$, we acquire

$$A + \log\left(f_2\left(a_K^*\right)\right) + \theta_K - \log\left(f_2\left(a_K^*\right)\right) - \tilde{\theta}_K = 0$$

That is,

$$A + \log\left(f_2\left(a_K^*\right)\right) + \theta_2 - \log\left(f_2\left(a_K^*\right)\right) - \left(\log\left(c_K\right) - \log\left(f_2'\left(\tilde{a}_K^*\right)\right) - \sigma_K^2/2\right) = 0$$

Thus then we obtain:

$$\log\left(f_2'\left(\tilde{a}_K^*\right)\right) = \log\left(c_K\right) - \theta_K - A - \sigma_K^2/2 = \log\left(f_2'\left(a_K^*\right)\right) - A \tag{10}$$

Since $A < 0$, $\log\left(f_2'\left(\tilde{a}_K^*\right)\right) > \log\left(f_2'\left(a_K^*\right)\right)$, we obtain $f_2'\left(\tilde{a}_2^*\right) \geq f_K'\left(a_K^*\right)$.

Due to the concavity of the functions f_1 and f_2, we obtain:

$$\tilde{a}_K^* < a_K^* \tag{11}$$

As for nitrogen fertilizer use, we have

$$f_1'\left(a_N^*\right)\exp\left(\theta_N + \sigma_N^2/2\right) = c_N.$$

$$\lambda f_1'\left({a_N'}^*\right)\exp\left(\theta_N + \sigma_N^2/2\right) = c_N.$$

Where ${a_N'}^*$ is the optimal nitrogen use under the misspecifed model, and a_N^* is the optimal nitrogen use in the true model. Obviously, we have ${a_N'}^* > a_N^*$ under the model of misspecified learning.

Proof of Prediction (2)

Prediction 2): If $a_N \to a_N^*$, $\tilde{a_K} \to a_K^* \Rightarrow \tilde{b}_K \to b_k$. If farmers are nudged to take the correct actions, then their subjective beliefs $\tilde{b}_k$ about the effectiveness of P/K will move toward the true state; then, undervaluation of P/K will decrease.

$$\begin{aligned}\tilde{b}_K\left(a_N', a_K', b_K\right) &= \log\left(C' + f_2\left(a_K'\right)\exp\left(b_K\right)\right) - \log\left(f_2\left(a_K'\right)\right) \\ &= log(\frac{(1-\lambda)f_1(a_N')exp(b_N)}{f_2(a_N')} + exp(b_K))\end{aligned} \tag{12}$$

We can clearly see that $\tilde{b}_K\left(a_N', a_K', b_K\right)$ is increasing as a_N' decreases and a_K' increases. Suppose $a_N'' > a_N' > a_N^*$, $a_K'' < a_K' < a_K^*$, then we'll have $\tilde{b}_K\left(a_N', a_K', b_K\right) > \tilde{b}_K\left(a_N'', a_K'', b_K\right)$.

where a_N' and a_N'' are different levels of nitrogen fertilizer input. a_K' and a_K'' are different levels of phosphorus/ potassium fertilizer applications.

Assumptions in Heidhues, Kőszegi, and Strack (2021)

In Heidhues, Kőszegi, and Strack (2021), they make three weak assumptions that bound the agent's misinference and the sensitivity of her misinference and behavior. We rewrite those assumptions as follow,

(i) There exists a constant $\Delta > 0$ such that $|b_K - \tilde{b}_K(b_k, a_N, a_K)| \leq \Delta$ for all b_K,a_N and a_K.

(ii) The function $\tilde{b}^K_{a^K}$ is bounded.

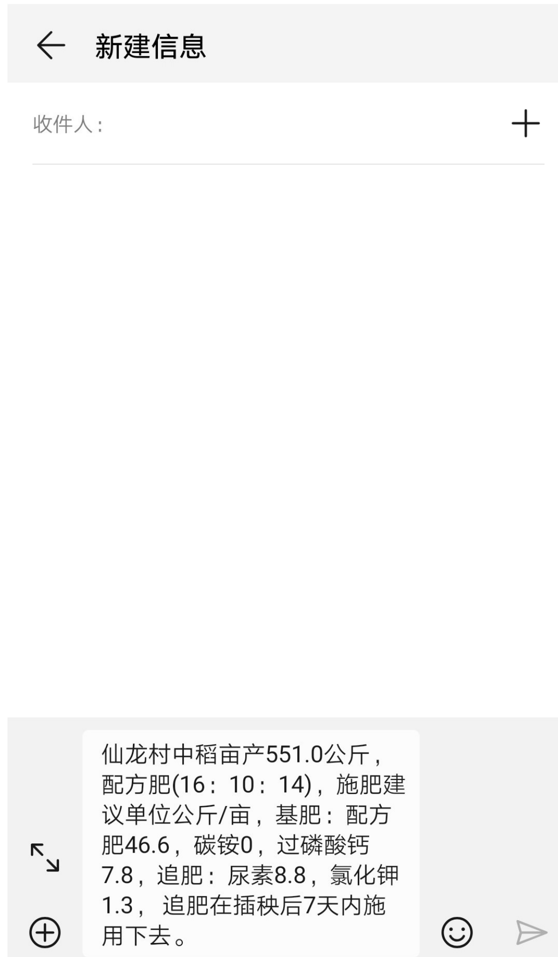
(iii) There exist constants $d, m > 0$ such that for any t and any $\tilde{\theta}_K$, we have $\left| a^*_K(t, \tilde{\theta}_K) - a^*_K(\tilde{\theta}_K) \right| \leq \frac{1}{t^m} d$.

Appendix C: Documentations about the Implement of Experiments

Figure C1: The Last Two Pages of the Mobile Application

(a) Connect Fertilizer Producer for Customized Fertilizer Recommendations

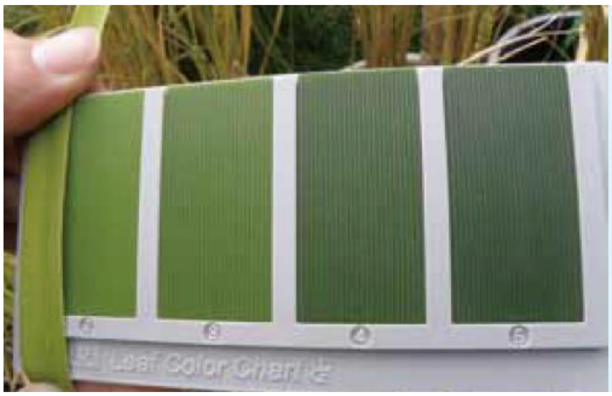
(b) Generate Recommendation Messages to Non-smart Phone

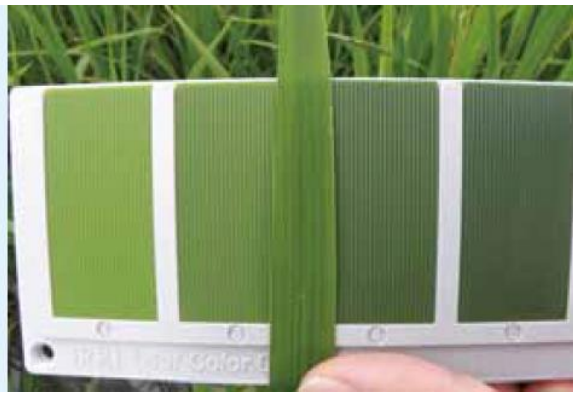


Note: Figure C1a shows that farmers can choose a specific fertilizer producer to order customized fertilizers by inputting names, contacts and coordinates of the plots. Figure C1b shows that the app can generate a message with fertilizer recommendations to non-smart phone users.

Use Leaf Color Chart to Help Fertilizer Application

This leaf color chart divides the greenness of rice crop's leaf color into six standard degrees. **Farmers can use the leaf color chart to diagnose the nutrition status of rice.** It helps diagnose the nitrogen concentration in rice crops and determine the optimal amount of **top-dressing nitrogen fertilizer usage**. When the amount of nitrogen fertilizer application is low, the number of tillers in crops will be low, and the color of leaves will be yellow. Conversely, when the amount of nitrogen fertilizer use is too high, the plant will have more tillers and the color of rice leaves will be close to dark green.





Instructions

1. **Randomly select more than 10 rice plants** that are not infected with the disease on their leaves in the rice field and take the leaf color during the rice **tillering period (growing stage) (20-30 days after transplanting)** and the **panicle-primordium-differentiation stage (40-50 days after transplanting)**, that is, before the second top dressing and panicle fertilizer application.

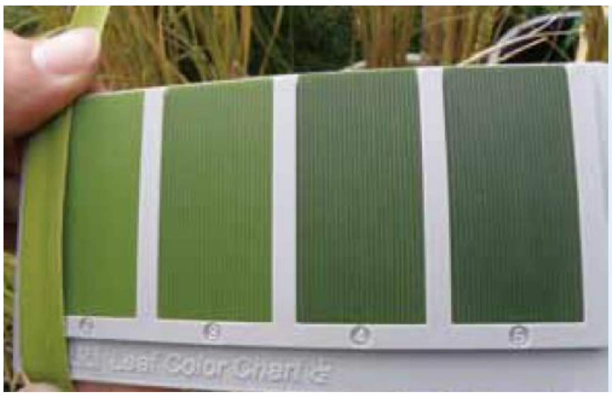
2. For each rice plant, **the middle section of the uppermost leaf which has been fully expanded** is selected to measure the leaf color.

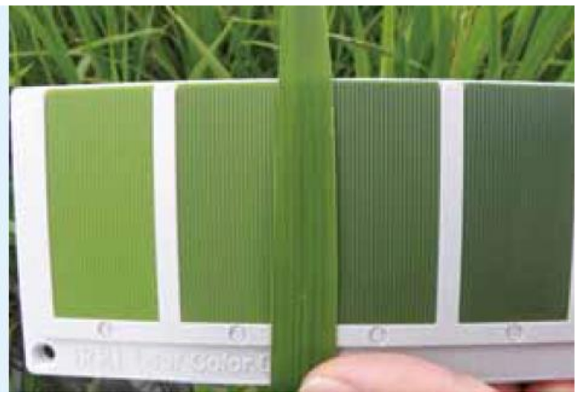
3. **First pull the leaves** whose leaf color is to be measured to **the inner shade of the investigator** to determine the leaf color, and compare the level of the leaf color (the leaf color board is divided into 1-6 levels according to the intensity). Avoid placing the leaf color board under the sun for direct observation, so as not to affect the accuracy of leaf color interpretation due to light reflection.

4. Determine the overall nutritional status of the rice in the area based on the average leaf color grade performance of 10 rice plants measured in the field, in order to determine the amount of topdressing nitrogen fertilizer used. **The optimum leaf color** ranges from **leaf color 3** to **leaf color 4**. When the leaf color is **lower than level 3** (the color is yellower than level 3), the rice plant is in **poor nutrition**, and **the amount of nitrogen fertilizer (urea) should be increased**. When the leaf level is **located at 3.5** (leaf color plate is between level 3 and 4), the leaf color is normal and the rice plant is in **good nutrition**, the amount of fertilizer can be used as normal. When the leaf color grade is at level 4, it indicates that the nitrogen content of the plant is slightly higher, and the nitrogen fertilizer should be less than the usual amount. When the leaf color grade is at **level 5** (leaf color is dark green), represents **the fertilizer application rate is too high, no more nitrogen fertilizer is needed**.

应用叶色板帮助水稻施肥

此叶色卡为 6 色叶色卡，将水稻叶色卡划分成六个标准色。**农户利用水稻营养状况诊断叶色卡，可直接比色读出水稻色级数。**来诊断水稻的含氮量，判断**追肥的**用量。氮肥施用量低时，植株分蘖数少，且稻叶色泽偏黄；相反的，当氮肥用量过高时，植株分蘖旺盛且稻叶色泽浓绿。





使用方法

1. 在水稻**分蘖盛期（插秧后 20-30 天）**和**幼穗分化期（插秧后 40-50 天）**，即第二次追肥及穗肥施用前，于水稻田间**随机选取 10 株以上**稻叶上未染病的稻株量取叶色。

2. 每株水稻选取**最上位已完全展开的叶片的中段部分**测量叶色。

3. 将欲测量叶色的叶片**先拉向调查者内侧**遮荫处进行叶色判读，并比对其叶色属于何种级距（叶色板上依叶色浓淡划分为 1-6 级）。避免将叶色板置于阳光下直接观测，以免因光线反射而影响叶色判读的准确性。

4. 依据田间量测 10 株水稻之平均叶色级距表现，判定该区水稻整体营养状态，以决定追肥氮肥使用数量。**最优水稻叶色**范围为 **叶色 3** 到**叶色** 4 之间**。**当叶色**低于 3 级**时（叶色较第 3 级黄），显示稻株营养**状况不良，应增氮肥（尿素）用量**。当叶色等级**位于 3.5 时**（叶色板等级介于第 3 级至第 4 级间），叶色属于正常值，显示**稻株营养状况良好**，可依惯行用量施肥。当叶色级数位于 4 级时，即显示植株氮素含量稍高，肥料施用应较惯行用量酌减氮肥。当叶色级**数达 5 级时**（叶色呈现浓绿色泽），表示肥料**施用量已过多，不需再施氮肥**。