**Title**
Three Essays on Development in China



**Author**
Lin, Wei

**Publication Date**
2022

Peer reviewed|Thesis/dissertation

# Three Essays on Development in China

by

Wei Lin

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Marco Gonzalez-Navarro, Chair
Professor Benjamin Faber
Professor Jeremy Magruder

Spring 2022

Three Essays on Development in China

Abstract

Three Essays on Development in China

by

Wei Lin

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Marco Gonzalez-Navarro, Chair

This dissertation studies two important fields of development economics in China, including agriculture (chapters 1 and 3) and education (chapter 2). Guided by theories in learning and decision-making literature, I exploit randomized controlled trial (RCT), lab experiment, and natural experiments to causally estimate the impact of learning (failure) on the productivity-related outcomes.

In Chapter 1 (coauthored with Binkai Chen and Ao Wang), we investigate 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.

Chapter 2 (coauthored with Binkai Chen and Ao Wang) intends to investigate the causal impact of collegiate economics courses on individual learning and decision-making under a development context. By exploiting a Chinese college-admission system that quasi-randomly assigns students to economics/business majors given students' preferences and the College Entrance Exam's cutoff scores for economics/business majors, we are able to isolate the treatment effects of an economics education on students' responses to a decision-making survey. Specifically, we compare the survey responses of students who narrowly meet the cutoffs for the economics/business majors to those who do not and find that students educated in economics/business courses are more likely to be risk neutral and less prone to common biases in probabilistic beliefs. While students in economics/business majors do not show significant changes in social preferences, they appear more inclined to believe that others behave selfishly.

Chapter 3 is joint with Qianmiao Chen and Shaoda Wang. We investigate how initial land endowments affect household short/long-term decision-makings on labor allocation when land and labor markets are incomplete. The village-level and family size-based land redistribution scheme before 2003 in China exogenously allocated households with differentiate initial land endowment. Using a rich rural fixed point survey data which tracks roughly 20,000 households during 1986 - 2013, we compare the decision-making of households that had population change before and after the last redistribution, the timeline of which is barely predictable by households. The empirical findings suggest that most of households oversupplied labor in agriculture, which led to the marginal product of labor deviate the optimum. Households with fewer initial land endowments decided to input more intense labor in agriculture, which exacerbated the labor misallocation given pre-existing over-employment issues. After forming the land rental market, households with fewer land endowments per capita rented more land, though more land did not influence household agricultural labor allocation. We then examine the impact of labor market reform on alleviating over-supply of labor. The overall agricultural labor input was reduced by 50-70%.

The findings in this dissertation can deepen our understanding of agricultural development and education system in China and shed lights on the research on individual decision-making process in other fields of development economics and other countries.

# Contents

# List of Figures

# List of Tables

# Acknowledgments

Throughout the writing of this dissertation, I have received tremendously generous support and encouragement from so many smart and nice people.

I would first like to thank my advisor, Professor Marco Gonzalez-Navarro. Marco supported and influenced me in many dimensions. Marco has shown great patience and offered invaluable guidance to me for every project of mine. Many times in my third and fourth years, I brought vague ideas to his office, seeking his advice. He always encouraged me to persevere, since no project is perfect at the beginning. Also during these years, I was stuck writing my job market project and often felt anxious. Marco often reached out and helped me through my anxiety. He also advised me to apply for many sources of research funding for large experiments, teaching me how to write applications sentence by sentence.

My committee member Professor Jeremy Magruder influenced my thinking as well in formulating research questions and methodology. Jeremy mentored me from the first year of my PhD journey, encouraging me to conduct my own experimental project, and guiding me through my oral exam and dissertation. His insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

I also have benefitted a lot from my committee member, Professor Benjamin Faber. Prior to my graduate studies at Berkeley, I was lucky to know Ben and serve as his research assistant on one of his projects in China. Thanks to Ben's recommendation letter, I received an offer of admission from Berkeley ARE on the first day of the Chinese Lunar New Year in 2016. Ben always encouraged me to think about the big picture of "good research" and served as my orals exam and dissertation committee member. I really enjoyed all the delightful and helpful conversations I had with Ben over the years.

In addition to my committee, many other individuals have provided excellent and selfless guidance and support. Professor Adam Szeidl spent countless hours supporting me throughout the past four years. In addition to being exceptionally sharp on helping with my papers technical details, Adam always urged me to think about the theory behind each empirical exercise, which really reshaped how I think about empirical papers in important ways. When I decided to go on the job market, Adam guided me a lot to help me finalize my search list, and always had my best interest in mind. Professors Elisabeth Sadoulet and Alain de Janvry also had a huge influence on my PhD study and career. I read a number of their papers and books prior to starting my own research and choosing my fields of study. They have been a source of endless encouragement and advice since my first year at UC Berkeley.

I am also deeply indebted to the entire development economics community at UC Berkeley. My research has substantially benefited from my interactions with Professors Thibault Fally, Supreet Kaur, Ethan Ligon, Aprajit Mahajan, Ted Miguel, Brian Wright, and many others. Of course, many of my dear friends have also supported and inspired me throughout the whole duration of PhD life. I am deeply grateful to my coauthors, Ao Wang, Jing Cai, Binkai Chen, Qianmiao Chen, Dennis Egger, Shaoda Wang, Xiao Ma, Lizhi Liu, Xiaojun Li, and Robert Gulotty. I have learned a substantial amount from their intelligent and insightful views. I really look forward to conducting more fruitful collaborations with them in the future. I also thank my peers at Berkeley for their emotional support throughout my PhD journey. The list includes, but is not limited to, Daniel Kannell, Shelley He, Yue Huang, Dong Wei, Robert Pickmans, Chaoran Yu, Zhimin Li, Wenfeng Qiu, Carly Trachtman, Jay Sayre, Leopold Biardeau, Yuen Ho, and David Qihang Wu. I am especially grateful to Danny being such a great foreign friend, and for all the computer games that we played together and the nice food that we tried together.

Funding for my projects was generously provided by CEGA, the Weiss Family Fund, the National Social Science Foundation of China, and the Special Fund from Central University of Finance and Economics.

# Chapter 1

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

## 1.1 Introduction

Existing theoretical and empirical literature identifies and explains learning failures in technology adoption in a single domain.[1] However, in reality agents often need to choose multiple technologies simultaneously. For example, agents face a trade-off problem between using antibiotics or improving hygiene conditions. Farmers face a trade-off between using pesticides or introducing bug-resistant crops. School managers must choose between building better facilities and improving teacher quality. Firm owners can either choose to increase managers' hours of work or improve management practices. In all these examples, the first technology is more immediately noticeable than the second, which causes agents to overuse the first, which crowds out the adoption of the second less-salient technology. Hanna, Mullainathan, and Schwartzstein (2014) begin to explore the role of mislearning in modeling the frictions caused by incorrect technology use. Yet little is known 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 under-

[1] For instance, agents may overuse a single technology 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 poor education (Cui et al.,

use and overuse of fertilizers are especially puzzling because farmers tend to employ these technologies over decades, which should be a sufficient time frame for them to learn their 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 primarily by activities in Brazil, Europe, India, the United States, 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 ($T_1$);[4] whether they received the soil analysis and customized and dynamic fertilizer recommendations through a smart mobile application ($T_2$); whether they received the soil analysis, the mobile application, and a training session from agricultural extension specialists ($T_3$); or whether they served as a pure comparison group. In the $T_3$ 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.

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 their use of phosphorus and potassium fertilizer relative to the control group. Specifically, we find that $T_2$ (App) and $T_3$ (App + in-person training) effectively reduced the use of nitrogen fertilizer 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 the use of nitrogen fertilizer occurred mainly during the growing stage of crops.

The overall use of phosphorus in the $T_2/T_3$ 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

---

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.

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

by 23.2-23.9 percentage points ($p < 0.01$), compared with a control mean of 3%. Similarly, for potassium we find that use increased by 1.37/2.89 kg per mu in the $T_2/T_3$ 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 and potassium, since they seldom applied these fertilizers prior to the experiment. In contrast, there were no significant effects on fertilizer use in the $T_1$ 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 subsequently explore the changes in yields/profits due to the optimization of different fertilizers. The change in fertilizer use caused by $T_2$ and $T_3$ led to a significant 5-7% increase in yields and a 6.0-6.9% increase in revenues, 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 in 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 by 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. The gap between their subjective beliefs and their 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, and such distorted actions influence their valuation of the effectiveness of the second technology. Consequently, agents overuse the first and under use the second. 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 composition and 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. 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 show that farmers' evaluation of the effectiveness of nitrogen remained unchanged since most farmers (95.2% in the control group) already understood its effects on greenness. We then explore the $T_2$ intervention, which encouraged farmers to change their fertilizer applications, but didn't involve direct contact to change their beliefs. We find that 22.0, 20.8, and 17.1 percentage points more farmers correctly understood the relationship between P/P/K and flower timing/root length/grain density after the $T_2$ 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 (LCCs) to help them better calibrate the optimal level of nitrogen to use 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 with the greenness on the charts to make informed decisions about optimal top-dressing fertilizer applications.

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 LCC intervention immediately reduced farmers' application of nitrogen fertilizer by 3.76 kg per mu—a 12.3% decrease compared with 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. 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 with the control group.

We next 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 data on fertilizer use 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 the data from 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 did not drive the overuse of nitrogen fertilizers. First, the farmers' bias with regard to fertilizer application focuses on the growing stages. Second, their decision-making is based mainly on greenness signals during the growing season. They believe the greener the better.

Taken together, the new mechanisms proposed by this paper are externally relevant in other contexts where agents learn about multiple technologies simultaneously.[7] Our re-

[6] In $T_2$, we provided only the mobile application and instructions to farmers; thus, the updating of beliefs should have come only from self-learning.

[7] For example, farmers face a trade-off between pesticides and bug-repellent crop varieties. Agents face a trade-off between taking antibiotics and improving hygiene. In these examples, the first technology generates more observable feedback, while the second technology has less salient impacts.

sults also demonstrate that cost-effective interventions guided by theory can correct agents' sub-optimal input choices, which has 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. 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 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 fall 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 the literature. Our research questions are most related to topics on technological learning and misuse of a single technology. Existing theories and empirical evidence in the field explore under-investment in agricultural technologies due to labor costs (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 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. To the best of our knowledge, our paper 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 can indeed resolve learning failure in fertilizer application.

Our interventions and policy implications leverage information communication tech-

[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*a*), 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).

[9]We compare our impacts with 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%).

nology (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 the economics literature but has a significant impact on the global environment.

The paper proceeds as follows. Section 1.2 describes the setting and first-phase experimental design. Section 1.3 discusses results from the first-phase field experiment. Section 1.4 presents our theoretical framework. Section 1.5 details the second-phase experimental design and interprets the results. Section 1.6 concludes.

## 1.2 Setting and First-Phase Experimental Design

### Fertilizer Consumption: China vs. the World

Fertilizer has been able to generate high returns for farmers and was responsible for significant growth in agricultural yields during the 20th century. In 2020, the global fertilizer market amounted to more than US$ 171 billion. With the considerable increase in fertilizer use, traditional agricultural extension programs have 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 smallholders may be ill-informed regarding safe and sustainable fertilizer use due to a lack of extension support. China provides a perfect context for studying the inappropriate use of fertilizer technology, since 98% of farming households have a farming plot of less than 2 hectares (Wu et al., 2018).

There is a longstanding discussion on fertilizer misuse in China.[10] As the world's largest fertilizer producer and consumer, China accounts for roughly 28.8% of global fertilizer use, while its arable land is only 7.6-9% of the world's total. Figure 1.1 shows the intensity of fertilizer application indexed by kilograms per hectare from 1961 to 2014

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

among some developed and developing countries, including Brazil, Bangladesh, China, Germany, India, Kenya, Mexico, and the United States. Starting in 1961, China's rate of fertilizer application grew rapidly, particularly after 1980 when synthetic fertilizers were introduced. Developing countries like Bangladesh, Brazil, and India, followed a similar growth pattern in fertilizer intensity as China in the earlier stage. For developed countries like Germany, fertilizer use reached a peak around 1980, and dropped dramatically after 2000,[11] when the German government started to impose restrictions limiting fertilizer application. From this cross-country comparison, we believe that it is important for the other countries to avoid following the same path of excessive fertilizer application.

Using a rich panel dataset that traces roughly 20,000 farming households from 1986 to 2015 over 360 villages, we depict the relationship between fertilizer application rate per hectare and farm scale. In Figure 1.2a, the top red line indicates the suggested amount of fertilizer application per hectare in western Kenya (Duflo, Kremer, and Robinson, 2008), roughly 242 kg/ha, and the bottom line shows the suggested level in India (Foster and Rosenzweig, 2010), roughly 160 kg/ha. Figure 1.2a highlights two important features. First, Chinese farmers tend to systematically overuse fertilizer at the intensive margin, which is consistent with the macro statistics that fertilizer intensity in China is four times the world average. Second, the downward sloping curve implies that smallholders have even higher application rates than farmers with large landholdings, suggesting poorer decision-making on fertilizer input by smallholders.

While underresearched in the economics literature, excessive use of fertilizers, especially nitrogen fertilizer, has resulted in massive emissions of greenhouse gases ($CO_2$, $N_2O$). As shown in Figure A.1 in Appendix A1, developed countries such as the United States, Australia, Germany, the Netherlands, and Denmark, and developing countries including India, Bangladesh,[12] and China are experiencing the same environmental concern regarding nitrogen overuse. Therefore, finding an appropriate way to improve fertilizer application can resolve the global concern for sustainable development and the fight against global warming (Tian et al., 2020).

## 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 determine whether farmers in China apply a sub-optimal mix of fertilizers and to understand the economic consequences of such misuse, we designed and implemented two experiments consisting of two main phases: (i) change farmers' actions in choosing the level of fertilizer application by providing individual soil analysis and customized fertilizer recommendations; (ii) correct farmers' overestimation of the return to nitrogen and greenness by distributing leaf color charts among 1,200 farmers in 200 villages.

[11] We ignore the period around 1990 since statistics might change due to reunification.

[12] See Brainerd and Menon (2014); Rahman and Zhang (2018); Islam and Beg (2021).

**Universal Soil Testing Program.** To provide farmers with customized recommendations, we first acquired the administrative data on the soil analysis from the universal soil testing program implemented by Hunan province. Figure 1.3a shows the distribution of universal soil testing program 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) soil texture (type of soil under soil taxonomy), pH levels, levels of primary macronutrients (nitrogen, phosphorus, and potassium), secondary macronutrients (calcium, magnesium, and sulfur), and level of organic matter.[13]

Our baseline survey, which was implemented in April 2020, shows that dissemination of the testing results to farmers was quite low. Less than 5% of 1,200 surveyed households reported that they had access to this information. We find two main obstacles that prevented information dissemination: 1) the soil testing results are too abstract and barely readable on farmers' side; and 2) the cost of face-to-face dissemination is high and the agricultural extension workers didn't have strong incentives to distribute it. Furthermore, in our surveys, less than 20% of farmers had a talk or received guidance from the agricultural extension workers on agricultural practices. Therefore, self-learning plays an important role in the application of different fertilizers.

**Recommendations.** To address obstacle (1) aforementioned, in addition to the soil analysis, we also randomly offer corresponding fertilizer recommendations.[14] The customized recommendations of fertilizer dosage that mixes N/P/K could be generated based on three elements: a reliable production function of different fertilizers, the crop-specific demand models for the micronutrients, and seasonal prices to maximize farmers' profits. The primary production function is simulated from 10,000 experimental trials in Hunan province, in which three fertilizers (nitrogen, phosphorus and potassium) were randomly adopted 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 crop-specific demand for the micronutrients is presented as a function of the micronutrients in the soil and targeted yields.

**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

[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

company,[16] co-developing a mobile application that has the following appealing features: 1) It is fully endorsed by Hunan and Leiyang governments (Figure A.2a) and can provide guidance for up to 15 crops (Figure A.2b). It connects the farmer to the administrative dataset on the universal soil testing for millions of local plots in Hunan province. Through GPS tracking (Figure A.3a) or by selecting the region (Figure A.3b), farmers can acquire the soil analysis results of the nearest testing plot where a sample of soil was collected and analyzed. 2) This smart mobile application displays the 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 A.4a. 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) together in the multiple cropping stages. As shown in Figure A.4b in Appendix A1, 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]

## Experimental Timeline and Framework

**Randomization.** We first draw a full list of 348 villages from the Department of Agriculture in local areas, with basic information including the total land area, average pH value, average N/P/K concentration in the soil, and organic matter in the plots. In the screening of villages with sufficient counts of rice farmers, we kept villages whose agricultural land area are more than 1000 mu (equivalent to 66.7 hectares). We listed these villages by the alphabetical order and then randomly selected 200 villages with the random number generator. Afterwards, as shown in Figure 1.4, we randomized these 200 villages into four arms ($T_1$, $T_2$, $T_3$ and Control). In each village, we randomly selected six rice farming households from the resident list provided by the 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 A3. To ensure farmers or their household mem-

[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.

[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.

bers 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 intervention builds on Hanna, Mullainathan, and Schwartzstein (2014), where they presented farmers with a summary of the trials' findings that pod size is important to the yields, and found the evidence of improving farmers' learning.

**4) Control group:** 300 farming households in 50 villages only received surveys.

We list the timeline of the data collection activities and the implementation of the first-phase experiment as follows,[18]

| | |
|---|---|
| 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 ($T_1$, $T_2$, $T_3$ and Control) in late March 2020, when China's economy was fully reopened.[19] The data

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

[19] China's economy was gradually reopened in March 2020 after coronavirus shutdowns and recovered

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 including (i) the input and output like yields, profit, land area, and other variables; (ii) the purchase and usage of different fertilizers in multiple stages of the cropping cycle; (iii) the testing results on individual soil quality from the nearest testing plot, distance between the farmer's plot to the testing plot, fertilizer recommendations predicted by the soil testing, and the gaps in fertilizer usage 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) fertilizers.

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 in Leiyang, Hunan province, helping them prepare for the incoming interventions in the $T_3$ group. In two weeks before the survey, we started to provide comprehensive training to the enumerators, who were recruited from local colleges and could speak local dialect. During the survey itself, we also sent several independent monitor teams to proctor the interview process to control the quality of questionnaire.

## Data and Balance Test

Table 1.1 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, the average yield was about 460-470 kilograms per mu, with the revenue (per mu of land) of 1096-1120 RMB. On average, for each mu of land, 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 in the growing stage. The application intensity for phosphorus and potassium fertilizers, on average, is only 0.84/2.16 kilograms per mu.

Panel B presents the soil testing results and predicted fertilizer recommendations. We are able to link the vast majority of farmers 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,[20] and the attrition. We present the statistics of two related questions

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.

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

as follows,

To measure farmers' beliefs about the return to greenness:

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

Among these options, 3) is well-documented in scientific and agronomic studies that the relationship between greenness and yield is inverted-U shape. Panel C shows that only a small proportion (7%) of farmers could give the correct response to this question.

To measure farmers' beliefs about the effectiveness of potassium, we ask the following question,

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

We mark in bold the correct answer. Survey data shows that most of the farmers (98%) didn't realize that potassium is effective in increasing the grains' density.

Overall, we do not find systematically significant difference between each of the treatment arms and control farmers in any of the outcomes, except for two variables that are statistically different at the 10% level. The first comes from the recommended N-P-K compound fertilizer between the $T_3$ group ($App + Visit$) and the control group. The second place of imbalance appears in the attrition rate between $T_2$ and control farmers. However, the attrition overall is quite low: roughly only 3-8 farmers in each treatment arm opt out of our study. Therefore, such imbalance should not be able to cause any major concerns.

## 1.3 First-phase Experiment Results

In this section, we establish several stylized facts that farmers simultaneously overuse nitrogen fertilizer and underuse phosphorus/potassium fertilizers. Our experiment shows that such sub-optimal input choice can be corrected by cost-effective interventions. We demonstrate that the provision of customized fertilizer recommendations through a mobile application improves farmers efficiency, and induces lower nitrogen application and increased phosphorus/potassium use. We then study mechanisms and design a theory-based intervention to test predictions of our model.

### Actual Use Deviates from Recommended Use

We begin the analysis with graphical evidence that highlights several main features in the data. Figure 1.5 presents the number of households (y-axis) against the extent of deviation (x-axis) in fertilizer application between farmers' actual use and the recommended.[21] The green bar and red bar suggest the simultaneous overuse of nitrogen and underuse of phosphorus and potassium, respectively. The degree of deviation in fertilizer use is present in a large domain, which indicates a 30-150% overuse for nitrogen fertilizer and 10-70% underuse for phosphorus/potassium fertilizers. The magnitude of the excessive use of nitrogen fertilizer is consistent with the conclusions in Wu et al. (2018), but we also find new evidence of phosphorus/potassium underuse.

In Figures A.5a, A.6a, and A.7a, we provide more detailed evidence about farmers' deviation in nitrogen fertilizer use across multiple cropping stages. Figure A.6a shows that the average usage of nitrogen fertilizer (blue line) overlaps with the mean recommendation (red line) in the planting stage, suggesting that in the planting stage, nitrogen fertilizer use barely deviated from the recommendations. Figure A.7a presents the fact that the overuse of nitrogen was mainly driven by farmers' practice in the growing stage (the used was much greater than the recommended). Turning to phosphorus fertilizer, we observe that the underuse (the used level was much lower than the recommended) appeared in both the planting stage (Figure A.9a in Appendix A1) and the growing stage (Figure A.10a in Appendix A1). The application of potassium fertilizers followed a similar pattern relative to phosphorus fertilizer: we observe that the underuse of phosphorus fertilizer existed in both the planting (Figure A.12a in Appendix A1) and growing stage (Figure A.13a in Appendix A1).

## Main Specification

To quantify the efficiency loss from simultaneous overuse and underuse of different fertilizers, we exploit the following regression specification,

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

where $Y_{iv}$ is the outcome of interest for household $i$ in village $v$ in the post-treatment period, including 1) the usage of different fertilizers; 2) the revenue, proifts, fertilizer cost and other input costs; 3) farmers' valuation of the effectiveness of different fertilizers. $T_1$, $T_2$, and $T_3$ are indicators for the corresponding treatment arms, which equal one if the village is assigned to the soil-testing group, mobile application group, and App plus training group, respectively. For inference, we cluster standard errors at the village level to reduce any correlated noise within the same village. Our coefficients of interest are $\beta_1$, $\beta_2$, and $\beta_3$, which measure the intention-to-treat effects of these three treatments.

## Results

Table 1.2 presents the regression results for the utilization of four different fertilizers (N, P, K, and N-P-K- compound fertilizers) as the outcomes. In columns (1)-(3), we compute the total use of individual nitrogen/phosphorus/potassium fertilizers by converting

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

compound fertilizer (use N:P:K = 15:15:15) into individual fertilizers.[22] We find that our recommendations were followed by farmers in the the $T_2$ and the $T_3$ groups. Staring with column (1) where the outcome is the total utilization of nitrogen fertilizer (after conversion). While the average intensity of nitrogen use of farming household in the control group was 30.84 kg/mu (equivalent to 462.6 kg per hectare), the usage in the $T_2$ (mobile application) group declined dramatically, by 3.92 kg/mu ($p < 0.05$), corresponding to a treatment effect of 12.17%. In addition, the usage of nitrogen fertilizer in the $T_3$ 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 usage of 14.74 kg/mu (equivalent to 221.1 kg per hectare) in the control group, the $T_2$ and $T_3$ interventions increased households' phosphorus usage by 2.34 kg/mu ($p < 0.01$) and 2.72 kg/mu ($p < 0.01$), respectively. Likewise, column (3) suggests that, the $T_2$ and $T_3$ interventions induced a higher potassium usage by 1.37 kg/mu ($p < 0.1$) and 2.89 kg/mu ($p < 0.01$), relative to the mean usage of 13.31 kg/mu (equivalent to 221.1 kg per hectare) in the control group. As could be expected, we do not detect any significant change in the adoption of nitrogen/phosphorus/potassium for farmers in the $T_1$ group, since it's hard for farmers to read and handle the soil analysis results by themselves.

The remaining columns demonstrates the treatment effects at the extensive and intensive margin across different cropping stages. Column (4) presents the treatment effects in the planting stage, showing that the use of compound fertilizer was increased in the $T_2$ and $T_3$ groups, which contributed to the reduction of nitrogen fertilizer reported in column (1) and the increase of total phosphorus/potassium fertilizers reported in column (2)/(3). Columns (5)-(9) present the treatment effects on the adoption of top-dressing phosphorus/potassium in the growing stage. While columns (6) and (8) report the rise of top-dressing phosphorus/potassium fertilizers per unit of land, columns (7) and (9) focus on the extensive margin — the proportion of households that used phosphorus/potassium as the to-dressing fertilizers. Although only 3% of farmers applied top-dressing phosphorus fertilizer during the growing stage prior to the treatment, the $T_2$ and $T_3$ interventions significantly increased the share of using phosphorus fertilizer by 23.9 and 23.2 percentage points ($p < 0.01$). In parallel, the treatment effects in $T_2$ and $T_3$ are 24.1 and 32.4 percentage points ($p < 0.01$) for the proportion of farmers using potassium fertilizer, compared to the mean of 9% in the control group. Altogether we conclude that Table 1.2 shows that our interventions lead to reduced nitrogen application at the intensive margin and increased phosphorus/potassium use both at the intensive and extensive margins.

**Closing the yield gap.** Table 1.3 explores the agricultural outcomes as farmers re-

[22]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.

optimized the mixed of different fertilizers after our interventions. Columns (1) and (2) present significant treatment effects on the yields for the $T_2$ and $T_3$ groups. Addressing the issues of nitrogen overuse and phosphorus/potassium fertilizer underuse led to a 22.74 and 31.65 kg/mu increase in yields for households in the $T_2$ and $T_3$ groups, relative to an average yield of 465.6 kg/mu in the control group. Turing to column (2), the treatment effect corresponds to a 5.4% ($p < 0.1$) and 6.7% ($p < 0.05$) increases in yields for farmers in the $T_2$ and $T_3$ groups. Column (3) demonstrates that farmers in the $T_2$ and $T_3$ groups had a higher profit than those of the control group, as a result of increased revenues and unchanged costs. As shown in column (4), the total revenues of farmers in the $T_2$ and $T_3$ groups moved up by 68.57 Yuan/mu and 78.79 Yuan/mu, respectively, accounting for a 6% ($p < 0.1$) and 6.9% ($p < 0.05$) growth regarding the total revenues. While column (5) shows that fertilizer costs didn't experience significant change, column (6) suggests that the costs of other inputs, including labor input, energy consumption, irrigation, and insurance stayed at the same level.

In summary, we find that $T_2$ and $T_3$ interventions both effectively helped farmers re-optimize fertilizer inputs and improve yields and profits.[23] The treatment effects are not statistically different between $T_2$ and $T_3$.

## How Large Was the Inefficiency?

To capture the efficiency loss, we first compare our treatment effects to other related interventions related to fertilizer usages. 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 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[24] 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.

[23]We also summarize these results in Figures A.5b for nitrogen fertilizers, A.8c for phosphorus fertilizers, A.11c for potassium fertilizers and A.14b for yields.

[24]1.8 billion mu arable land in total for different kinds of crops in China.

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).

## 1.4 Theoretical Framework

It's really puzzling that fertilizers are underused and overused simultaneously since farmers have been working with these technologies over decades, which should be a sufficient time frame for them to correctly learn the return of these fertilizers, according to standard learning-by-doing models. In addition, as rational farmers, they should be able to carry out own experimentation to find the optimal amount of different mix of fertilizers since land and fertilizers are are divisible. Despite the availability, the prevalence of misuse suggests that there is clear cognitive barrier in learning and belief formation that prevents them from choosing the optimal application. As such, understanding what exactly the cognitive barrier is and to what extent farmers can benefit from overcoming could help the farmers go beyond merely using the technology, but to use it efficiently. In this section, we explore the mechanisms behind such misuse of technologies.

### Stylized Facts and Survey Evidence

Why did farmers overuse nitrogen but underuse phosphorus and potassium fertilizers? To answer this question, we dig deeply into the functions of different fertilizers. Fertilizers contain three vital dimensions —nitrogen (N), phosphorus (P), and potassium (K) — among which nitrogen (N) produces salient signals by increasing the greenness of the crops in the growing stage, while the effects of phosphorus (P) and potassium (K) do not generate salient signals. Specifically, phosphorus (P) can mainly boost the root length and change the timing of flowering, while potassium (K) could enhance the density of the rice grains, which are barely observable.

We presents the recommended practice suggested by the specialists for optimal fertilizer application in different cropping stages in Figure 1.6a. The cropping cycle can be divided into three stages, as shown from left to right: 1) transplanting/planting stage, in which the basal fertilizer is applied; 2) growing stage, in which the top-dressing fertilizers are strongly recommended; and 3) ripening and harvest stage, in which no fertilizer is needed. In the transplanting stage, farmers are advised to apply N-P-K compound fertilizer as the basal fertilizer, while the extra top-dressing nitrogen and phosphorus/potassium fertilizers are recommended to be used during the growing stage. Meanwhile, farmers receive the signals through the greenness of the crops in the growing stage.

Figure 1.6b shows farmers' actual fertilizer application in the baseline, with the share of households that adopted different fertilizers on the y-axis and the timing on the x-axis. Consistent with the recommended practice, the red bar suggests that, almost all the 1,200 farmers applied N-P-K compound fertilizer as the basal fertilizer in the planting stage. However, the green, orange, and blue bars demonstrate the inconsistency between farmers' actual use and the recommended practice in the growing stage. We observe that farmers only applied and adjusted nitrate fertilizer (N) during the growing stage (after the second week), but did not add any phosphorus (P) and potassium (K) fertilizers in the same period. Combining with the findings presented in Table 1.2 that the overuse of nitrogen mainly came from the growing stage, we hypothesize that, because farmers can observe the greenness signals which reflect the effectiveness of nitrogen during the growing stage, they will only adjust the usage of nitrogen fertilizer accordingly. Since they cannot observe any salient feedback from applying phosphorus (P) and potassium (K) fertilizers, they cannot learn the effectiveness of these two fertilizers and thus then underuse it.

To verify this puzzle, we further elicit farmers' beliefs about the relationship between yield and greenness. The question is presented 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 1.7 plots farmers' beliefs about the relationship between greenness and yields. Most farmers believed that the greenness is always positively correlated with yields (option 1). Only a small proportion of farmers (less than 7%) understood the true production function between greenness and yields — an inverted U-shape relationship (option 3), which is well documented in agriculture literature. Given this incorrect prior and the salient signals from nitrogen, farmers persistently over-applied nitrogen because of their overestimation of the return to greenness (misspecified production function). We display the basic intuition in Figure 1.8. First, the salient feature of nitrogen induces farmers to feel overoptimistic about the return of nitrogen fertilizers. Such exogenous overestimation endogenously causes farmers to overuse nitrogen fertilizer and expect a great yield. At the end of each period, farmers observe that yields are lower than expected and they would rationalize such lower-than-expected yield to the low-then-true-perception of the effectiveness of phosphorus (P)/potassium (K) fertilizers, since they believe that nitrogen fertilizers are productive and effective. As a result, it further causes farmers to underuse of phosphorus (P)/potassium (K).

Based on these stylized facts, we leverage recent behavioral economics theory 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 the overuse, and then such distorted action in turn affects farmers' beliefs about the effectiveness of other technologies abd causes the underuse. We also demonstrate how farmers' actions and beliefs evolve over time and converge to a sub-optimal equilibrium.

## 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:

$$\Pi(a_{Nt}, b_{Nt}, a_{Kt}, b_{Kt}) = f_1(a_{Nt})\exp(b_{Nt}) - c_N a_{Nt} + f_2(a_{Kt})\exp(b_{Kt}) - c_K a_{Kt}, \tag{1.1}$$

Where the functions $f_1$ and $f_2$ are concave, $f_i > 0, \lim_{a\to\infty} f_i'(a) = 0$. $a_{Nt}, a_{Kt} > 0$ is the amount of N-fertilizer/PK-fertilizer that farmers use at $t$, $c_{Nt}$ and $c_{Kt}$ are the normalized unit costs of N-fertilizer and P/K-fertilizer, 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. $\exp(b_{Nt})$ and $\exp(b_{Kt})$ represent external states as a productivity shiftier, which influence the marginal product of $a_{Nt}$ and $a_{Kt}$, and are correlated across farmers and time. In this context, we can interpret $\exp(b_{Nt})$ and $\exp(b_{Kt})$ as the realized effectiveness of nitrogen (N) and phosphorus (P)/potassium (K) fertilizer.

We further 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 $\epsilon_{Nt}$ and $\epsilon_{Kt}$ are independently normally distributed random variables with mean zero and variance $\sigma_N^2, \sigma_K^2$. $\Theta_N$ and $\Theta_K$ can be interpreted as **the average effectiveness** of nitrogen fertilizer and phosphorus/potassium fertilizers. The farmer need to learn about these two parameters as a

Bayesian learner. But $b_{Nt}$, which measures the effectiveness of nitrogen fertilizer, can also be observed by the farmer through greenness.

**Farmers' subjective beliefs.** Farmers' prior is that $\Theta_N$ and $\Theta_K$ are normally distributed with means $\theta_{N0}$ and $\theta_{K0}$, and with variances $v_{N0}$ and $v_{K0}$, respectively. Motivated by the stylized facts in Section 1.4, while the farmer understands the general form of the profit function, she misunderstands the specification of production function of nitrogen fertilizer $f_1$ as follows:

$$\tilde{\Pi}\left(a_{Nt}, b_{Nt}, a_{Kt}, \tilde{b}_{Kt}\right) = \tilde{f}_1\left(a_{Nt}\right)\exp\left(b_{Nt}\right) - c_N a_{Nt} + f_2\left(a_{Kt}\right)\exp\left(\tilde{b}_{Kt}\right) - c_K a_{Kt}, \tag{1.2}$$

where $\tilde{f}_1(\cdot)$ indicates the exogenous misspecification in the production function of nitrogen fertilizers. $\tilde{b}_{Kt}$ captures farmers' misunderstanding in their subjective beliefs about the effectiveness of phosphorus and potassium fertilizers, which deviates the true effectiveness $b_{Kt}$.

To be specific and without loss of generality, we define $\tilde{f}_1(\cdot) = \lambda\tilde{f}_1(\cdot)$, and acquire the profit function as:

$$\tilde{\Pi}\left(a_{Nt}, b_{Nt}, a_{Kt}, \tilde{b}_{Kt}\right) = \lambda f_1\left(a_{Nt}\right)\exp\left(b_{Nt}\right) - c_N a_{Nt} + f_2\left(a_{Kt}\right)\exp\left(\tilde{b}_{Kt}\right) - c_K a_{Kt}, \tag{1.3}$$

where $\lambda$ represents the degree of misspecification in the production function due to her misperception about the return to nitrogen. 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 $\tilde{\Pi}\left(a_{Nt}, b_{Nt}, a_{Kt}, \tilde{b}_{Kt}\right)$.

To solve for the equilibrium, we then impose three assumptions, which are motivated by the stylized facts aforementioned.

**Assumption 1.** $\lambda > 1$. *The farmer's degree of misspecification in the production function is greater then one.*

Consistent with our survey evidence in Figure 1.7, this assumption is the core exogenous misunderstanding of the farmer the she overestimates the return of nitrogen fertilizers. One possible explanation is that the use of nitrogen fertilizer has increased across time as shown in Figure 1.1, thus it is plausible that farmers would have formed the belief that more greenness leads to greater yields since they have spent most of their time on and observed the upward-sloping part of the curve. This argument is also endorsed by the mechanism that "what you see is all there is" by Enke (2020).

**Assumption 2.** *The farmer may misunderstand the effectiveness of P/K-fertilizers $b_{Kt}$ as $\tilde{b}_{Kt}$, since the effects of P/K-fertilizers are barely observable during the growing stage.*

For their beliefs about the effectiveness of N-fertilizers $b_{Nt}$, the farmer can correctly learn it in a certain period since nitrogen produces salient signals and timely feedback through greenness. In our sample, 93.8% of farmers understand that greenness is enhanced by nitrogen (N), while only 4.4% and 2.5% of farmers know that the effect of phosphorus (P) on the root development and the impact of potassium (K) on the grain's density.

**Assumption 3.** *For any input level $a_{Nt}$ $a_{Kt} \in \mathbb{R}$, and any level of effectiveness $b_{Nt}, b_{Kt} \in \mathbb{R}$, there exists a farmer's subjective belief $\tilde{b}_{Kt}$ such that farmer's perceived profits equal the realized profits, $\tilde{\Pi}(a_N, a_K, b_N, \tilde{b}_K) = \Pi(a_N, a_K, b_N, b_K)$.*

Based 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. We also impose some weak technical assumptions. Please see detailed proofs in Appendix A2 for reference.

**Belief updating.** The farmer chooses input level of these two fertilizers ($a_N$ and $a_K$) in each period to maximize her perceived expected profits ($\tilde{Q}$) in that period. As the farmer's priors ($\Theta_N$, $\Theta_K$) 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.

For the dimension of nitrogen, we assume that farmers can learning its effectiveness correctly through the observation of greenness. Another stylized fact noteworthy is that Chinese farmers started to use nitrogen fertilizer in 1940s, but only massively used phosphorus (P)/potassium (K) in 1990s. Such a long time duration between 1940-1990 should be sufficient for farmers to learn the effectiveness of nitrogen. So at the timing when farmers started to learning phosphorus (P)/potassium (K), we could treat $b_N$ as a constant and true value.

Therefore, the core learning is related to the effectiveness of phosphorus (P)/potassium (K).

**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:

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

Where $\tilde{\theta}_{K0}$ is the initial prior of the effectiveness. Since there is an exogenous misspecification in the profit function of nitrogen, this exogenous overestimation induces the farmer to mistakenly update her belief about the effectiveness of phosphorus (P)/potassium (K) as $\tilde{b}_{Kt}$.

## Predictions and Simulations:

After observing the profit at the end of period $t$ generated by the realized states $b_N$ and $b_K$, the farmer believes that the effectiveness of nitrogen fertilizer $b_{Nt}$ and phosphorus

(P)/potassium (K) fertilizer $\tilde{b}_{Kt}$ satisfying:

$$\tilde{\Pi}\left(a_{Nt}, a_{Kt}, b_{Nt}, \tilde{b}_{Kt}\right) = \pi_t = \Pi\left(a_{Nt}, a_{Kt}, b_{Nt}, b_{Kt}\right) \tag{1.4}$$

where $\pi_t$ is the realized profits observed by the farmer at the end of $t$. We then derive the subjective belief about the effectiveness of phosphorus/potassium $\tilde{b}_K$ as follows,

$$\tilde{b}_K\left(a_N, a_K, b_K\right) = \tilde{b}_K\left(a_K, b_K\right) = \log\left(C + f_2\left(a_K\right)\exp\left(b_K\right)\right) - \log\left(f_2\left(a_K\right)\right) \tag{1.5}$$

Where $C = (1-\lambda) f_1\left(a_N\right)\exp\left(b_N\right)$.

The interpretation for this equation is that the farmer believes that whatever action she chooses, she can infer a unique signal $\tilde{b}_K$ about $\Theta_K$ to equalize her belief and the real profit.

We then 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.** *The **exogenous** misspecification in the production of nitrogen induces farmers to overuse nitrogen fertilizer and **endogenously** undervalue the effectiveness of P/K fertilizers, and then under-investment in phosphorus (P)/potassium (K) fertilizers. Consequently, farmers will be trapped in a sub-optimal equilibrium with lower profits. They could become better off by re-optimizing the mix of different fertilizers.*

**Prediction 2.** *If farmers are nudged to adopt less nitrogen fertilizer and more phosphorus (P)/potassium (K) fertilizers in their actions in the current period, then their subjective beliefs $\tilde{b}_K$ about the effectiveness of P/K in the next period will move upwards toward the true value o the effectiveness. As such, farmers' undervaluation of P/K will decrease. See the details in Appendix A2 for the derivation of this prediction.*

**Prediction 3.** *Correcting farmers' overestimation/misspecification of the return to nitrogen fertilizer can lead to a lower nitrogen fertilizer application, and higher phosphorus (P)/potassium (K) fertilizer usage. Their belief about the effectiveness of phosphorus (P)/potassium (K) could also converge to the true value. To be specific, revising farmers' overestimation could reduce their nitrogen application immediately (within the same period), but also induce gradual learning about phosphorus and potassium fertilizers.*

Prediction (1) is verified by the point estimate in Tables 1.2 and 1.3. We present the simulations in Figure 1.9, showing the convergence of the beliefs about the effectiveness of nitrogen and phosphorus (P)/potassium (K) fertilizers. Consistent with our theoretical predictions and survey evidence, farmers' belief about the effectiveness of P/K slowly converges to an equilibrium which is below the true value, as shown in Figure 1.9c, while Figure 1.9b shows that the learning of nitrogen has a much faster speed.

## 1.5 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 misuse behavior is corrected; 2) design and implement a second-phase experiment to test model prediction (3); and 3) discuss the results from the second-phase experiment.

### Test of Predictions 2

Prediction 1) is tested in Tables 1.2 and 1.3, while, for prediction 2), we present the results in Table 1.4 using regression specification (1) in Section 1.3 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 the share 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 1.4, the outcome variables in columns (2)-(5) are binary variables which take the value of 1 if farmers gave the correct response to the corresponding questions. Column (2) shows that farmers' understandings of the effectiveness of nitrogen on greenness in the $T_1$, $T_2$, and $T_3$ villages are not significantly different from that of the control group, since 95% of farmers already understood it correctly. Column (3) presents the treatment effects of farmers' learning about the effectiveness of phosphorus on flowering timing in the growing stage. Relative to the average value of 13 percentage points in the control group, 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 effectiveness of phosphorus on root length, which shows the same direction 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 the share of 2.8% of farmers understanding the effect of potassium in the control group.

In summary, Table 1.4 provides direct evidence on model prediction (2) that farmers' undervaluation of the effectiveness of phosphorus/potassium will decrease if their underuse behaviors are corrected, but their 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 might be driven by social learning. In $T_2$ group, farmers can only figured out the effectiveness by themselves (individual learning). In the $T_3$ group, in addition to individual learning, farmers could also have new interactions with the agricultural extension agents, which can boost their learning (social learning).[25]

## Second-phase Intervention

To test model prediction 3) in Section 1.4, in this section, we design and implement a second-phase experiment by introducing the leaf color chart, aiming to correct farmers' overestimation of return to greenness (changing $\lambda$ to 1).

**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.[26] 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.[27] We also prepared a brochure with detailed instructions, showing farmers the correct timing, position, and method to use the leaf color charts. Our enumerators stayed in person with farmers and, on average, spent roughly 30 minutes in demonstrating the correct usage of the leaf color charts.

As shown in Figure 1.11, to improve farmers' learning and correct their overestimation, we distributed the leaf color charts among the 150 villages that were selected into the first-phase interventions ($T_1$, $T_2$, and $T_3$).[28] In Table 1.5, we conduct an additional balance test between farmers in $T_1$ group and the control group, and do not find any significant

[25]Prior to our interventions, only 20% of farmers had ever been instructed by the agricultural extension agents.

[26]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.

[27]See more details in Appendix A3, which demonstrates the full guidance and application process.

[28]Ideally, the cross-randomization could help us better identify the pure effect of the leaf color chart intervention. However, there are two concerns that prevent us from conducting cross-randomization. First, our partners, the local governments, hoped we could keep the experimental structure as simple as possible.

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 |

The second-phase intervention was implemented in December 2020, right after the first follow-up survey conducted in November 2020. In August 2021, we conducted a second follow-up survey 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 model and prediction (3) is true, then we should be able to see that the leaf color chart effectively corrects farmers' overestimation of the return to greenness, and simultaneously accelerates their learning about phosphorus and potassium in terms of usage.

## Results from the Second-phase Intervention

Table 1.6 reports the regression results using the regression specification in Section 1.3. We start the analysis with farmers in $T_1$, the coefficient in column (1) suggests that the leaf color chart intervention leads to an immediate drop in aggregate nitrogen usage by 3.76 kg per mu ($p < 0.05$), which is consistent with our model prediction (3) that learning nitrogen

---

Our survey activities cannot go smoothly without their coordination. 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.

is fast and salient. 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 adoption of top-dressing phosphorus and potassium fertilizers in the growing stage. Perhaps the large variation in the outcome variables results in such insignificance.

To further explores the potential effect of the leaf color chart on the adoption of phosphorus and potassium fertilizers, we decompose fertilizer applications into two stages: the planting stage and growing stage. Column (4) reports the application of N-P-K compound fertilizer which is used in the planting stage. Since the leaf color charts can be only used in the growing stage, we do not observe any significant change for compound fertilizer during the planting stage. However, we indeed find inspiring results in columns (7) and (9), which report the proportion of farmers using phosphorus and potassium fertilizers in the growing stage. The positive and significant coefficients in columns (7) and (9) suggest that, the leaf color chart intervention has led to a 6 percentage points increase in the proportion of farmers using top-dressing phosphorus and potassium fertilizers in the growing stage ($p < 0.01$ and $p < 0.1$).

As suggested in Hanna, Mullainathan, and Schwartzstein (2014), while the dimensions of technology are large, agents' attention is limited. An agent can only learn about the dimensions that she pays attention to. In our context, our treatment effects indicate that some farmers were starting to conduct their own experimentation with phosphorus and potassium fertilizers (columns (7) and (9)), perhaps because their binding attention was relaxed from nitrogen due to the help of the leaf color charts. Thus they were getting more confident about conducting experiments on phosphorus and potassium fertilizers. 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 via the leaf color charts not only immediately improved their learning of nitrogen dimension, but also simultaneously induced 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 the leaf color charts did not significantly enhance the pre-existing treatment effects by 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 fertilizer inputs in the first-phase experiment. In the meantime, the leaf color charts did not affect the existing treatment effects by 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 quite plausible that the leaf color chart intervention has no add-on effect.

Table 1.7 presents the effects of the leaf color chart intervention on farmers' perceptions of the return to greenness and their beliefs about the effectiveness of different fertilizers. 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. Consistent with the results in Table 1.4, farmers in the $T_2$ and $T_3$ groups, reduced their 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 $T_1$ group who only received the leaf color chart treatment, column (1) shows that their overestimation of the return to greenness declined, reflected by the point estimate that the share of farmers that chose the inverted U-shape relationship between greenness and yield increased by 27 percentage points ($p < 0.01$). This pure influence of the leaf color chart is very close to the effects of the customized fertilizer recommendations in the $T_2$ and $T_3$ groups (31.0 percentage points). While farmers had no change in their understanding of the effectiveness of nitrogen on greenness, since 95% of them already knew it, we also do not see any increase in their understanding of the effectiveness of phosphorus and potassium in the current period. It's because that farmers didn't update their beliefs and realize the effectiveness of phosphorus and potassium, since they had not yet observed the new yields and leaned under Bayesian rule.[29] We expect to see the belief updating after farmers produce and observe the new yields and profits in the next period, which inspires our future follow-up survey and research design.

Overall, compared to customized fertilizer recommendations, we find the second-phase intervention via the leaf color charts also effectively corrected farmers' overestimation of the return to greenness and improved learning of nitrogen/phosphorus/potassium. But such learning is sole depending on self-learning, which is developed in a gradual process and requires a longer learning period. All of these results are summarized in Figures A.5, A.6, A.7, as well as in Figures A.8, A.9, A.10, A.11, A.12 and A.13 in Appendix A1. Nitrogen fertilizers were overused in the growing stage only, while phosphorus and potassium fertilizers were underused throughout the cropping cycle (both planting stage and growing stage). From the perspective of policy design, policymakers face a trad off between choosing the customized recommendations and the leaf color chart interventions. The former treatment is more effective immediately but costly, while the latter is more scalable but less effectiveness. Policymakers should balance this trade-off based on budget and time constraints.

## 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 deviation in fertilizer use between the actual and recommended. In Table 1.8, columns (1), (2) and (3) present the treatment effects of the first-phase interventions on the difference between the actual application and recommended use. The outcome variables in columns (1), (2), and (3) are the nitrogen use gap [used - recommended], the

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

phosphorus use gap [used - recommended], and the potassium use gap [used - recommended], respectively. In the first phase, $T_2$ and $T_3$ dramatically reduced the nitrogen gap by 3.6 kg/mu and 4.7 kg/mu, corresponding to a 50% dip of the average gap in the control group. Similarly, the interventions also 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 of the leaf color chart intervention on closing the gaps of fertilizer application. Since the testing data were acquired through GPS tracking, in Table A.1 in Appendix A1 we show that the gaps are not statistically correlated with the distance between a farmer's plot and the nearest soil testing plot, in either the first-phase or second-phase experiemnt.

We then estimate the causal impact of the resulting loss in yields due to misuse of fertilizer by using the following IV strategy.

*(1) Sample Selection.* We limit the regression sample to those who simultaneously overuse nitrogen fertilizer and underuse phosphorus/potassium fertilizers so that the underlying relationship between fertilizer gaps and yields would be clearly defined. The overuse of nitrogen and underuse of phosphorus/potassium directly lead to lower yields. In Table A.4 and A.5 in appendix A1, we relax this restriction and the results are still consistent and robust.

*(2) Main Outcomes.* We calculate the gaps of each type of fertilizer by aggregating them across different stages, $N\_Gap$, $P\_Gap$, and $K\_Gap$ as the main outcomes, where $N\_Gap$=(total nitrogen used - recommended), $P\_Gap$=(total phosphorus used - recommended), and $K\_Gap$=(total potassium used - recommended).

*(3) Identification.* In Table 1.8, we show that $T_2$ (App) and $T_3$ (App + Training) interventions effectively reduced nitrogen/phosphorus/potassium gaps, suggesting that $T_2$ and $T_3$ can be served as valid instruments for the fertilizer gaps. But it also raises the issue of underidentification: it has 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.

The following equation (1.6) characterizes the causal effect of $T_2$ and $T_3$ interventions on $Gap^2$.

$$logGap_i^2 = \alpha' + \beta_1' T_{2i} + \beta_2' T_{3i} + \epsilon_i' \tag{1.6}$$

While equation (1.7) captures the effects of deviation in fertilizer application measured by $Gap^2$ on yield.

$$Yield_i = \lambda + \delta \widehat{logGap_i^2} + \epsilon_i'' \tag{1.7}$$

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 $\delta$ measures the effect of deviation in fertilizer application on yields.

Table 1.9 reports the point estimates of regressions (1.6) and (1.7). Column (1) presents the coefficients $\beta_1'$ and $\beta_2'$ in the first-stage regression, which again suggests that $T_2$ and $T_3$ interventions effectively reduced the natural log of $Gap^2$ by roughly one unit. The second-stage regression estimate 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) for robustness purpose, 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 make the magnitude more intuitive ans straightforward, our $T_1$ and $T_2$ interventions in the first phase reduced the outcome $Gap^2$ by 100% (column (1)), corresponding to a 7.7% increase in yields, which is consistent with our reduced form regression results of 7% in Table 1.3. To check the validity of using $T_2$ and $T_3$ as instrumental variables, in Table A.2 we replicate the IV-2sls regression using the baseline data and employing equations (1.6) and (1.7). We do not find any significance either in the first-stage or second-stage regressions since $T_2$ and $T_3$ did not affect the fertilizer applications and yields in the baseline, which further confirms the validity of the instrumental variables.[30]

**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 1.6, 3.76 from column (1), 0.48 from column (2), and 0.173 from column (3). Then we obtain $\Delta Gap^2$ as:

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

The naive estimation suggests that the leaf color charts reduced the log of $Gap^2$ by 0.44 points, indicating a 44% decline in $Gap^2$. This decrease corresponds to a underlying gain in yield by 3.4% or 16 kg/mu. In Table 1.10, we additionally present the effects of deviation in fertilizer application on revenues, fertilizer costs, and other costs using regression equations (1.6) and (1.7). Column (2) suggests that a ten percent drop in $Gap^2$ will increase revenues by 9.82 RMB per mu.[31] By plugging in the value of 44% decline in the $Gap^2$ , we estimate that farmers' 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 similar and robust when we employ the full sample from the second follow-up survey.

[30] In Table A.3, we conduct another type of IV regression. In columns (1), (2) and (3), we then estimate the effects of each gap on yields separately by instrumenting the gap in nitrogen use [Used - Recommended], gap in phosphorus use, and gap in potassium use separately with the $T_2$ and $T_3$ indicators. The results are still consistent with that of the IV regression presented in Table 1.9.

[31] We do not find significant impacts of $Gap^2$ on fertilizer costs and other input costs, which are consistent with previous findings presented in Section 1.3.

In summary, based on the back-of-the-envelope estimation, our second-phase intervention that provides leaf color charts led to an increase in yields by roughly 3.4% and revenues by 4%.

## Additional Discussions

In this section, we discuss some general applications, issues related to identification and interpretation, as well as some potential alternative explanations for nitrogen overuse.

**Applications in Other Fields.** Though this paper focuses on agricultural farming, our finding that the misperception of one technology could affect the valuation of the effectiveness of other technologies has more general implications. For example, in the filed of health studies in less-developed areas, there is a trade-off problem between taking antibiotics and improving hygiene conditions. Since the effects of antibiotics are immediate and can easily be observed, while the effects of improving hygiene conditions on health are less salient, we could see that antibiotics are often abused in the real life. Likewise, farmers are going to overuse pesticides if they face a trade-off problem between using pesticides and bug-resistant crops since pesticides generate more salient feedback. School administers are going to spend more money on building better facilities then improving teachers' qualities due to the salience of better facilities.

**Measurement Error in Self-Reporting.** A natural concern of our data is that our main outcomes of interest, like yields and inputs are self-reported, which raises the possibility that 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 are very likely to influence the usage of different fertilizers into different directions. Second, we mitigate this concern by collecting farmers' responses based on three different question modules regarding fertilizer use: 1) the total usage of different fertilizers; 2) the 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 1.12a; and 2) only a small proportion (2.43%) of them took the advice from fertilizer sellers seriously and followed the recommendations in Figure 1.12b. One plausible explanation is that sellers, unlike us, didn't have science-based knowledge on the soil quality of individual plots, and therefore they were able to just provided the average recommendations, which is less convincing. Besides, by correcting farmers' beliefs to induce them to use more phosphorus and potassium and less nitrogen, fertilizer traders 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 1.12c and 1.12d 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 from neighbors. The latter argument is consistent with the fact that fertilizer quality was high in our experimental site.[32]

**Price Difference and Budget Constraints.** 1) Farmers may use nitrogen more since it has a relatively lower price than the other two fertilizers. We directly elicit farmers' decision-making in terms of nitrogen usage. In Figure 1.13a, 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 1.13b). 2) Farmers use nitrogen more than other two fertilizers since it has a lower price than the other two fertilizers. Contradicting this explanation, we find farmers only deviated nitrogen fertilizer use during the growing stages when they received signals from crops (Figures A.5, A.6, and A.7), 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.

## 1.6 Conclusion

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 the overestimation of the return to greenness, which reflects the primary effect of nitrogen. In their subjective beliefs, the lower-than-expected output is rationalized by the lower-than-true perception of the effectiveness of phosphorus/potassium fertilizers. Our interventions, including the provision of customized fertilizer recommendations through a mobile application and leaf color charts, effectively reduced nitrogen fertilizer application, increased phospho-

[32]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.

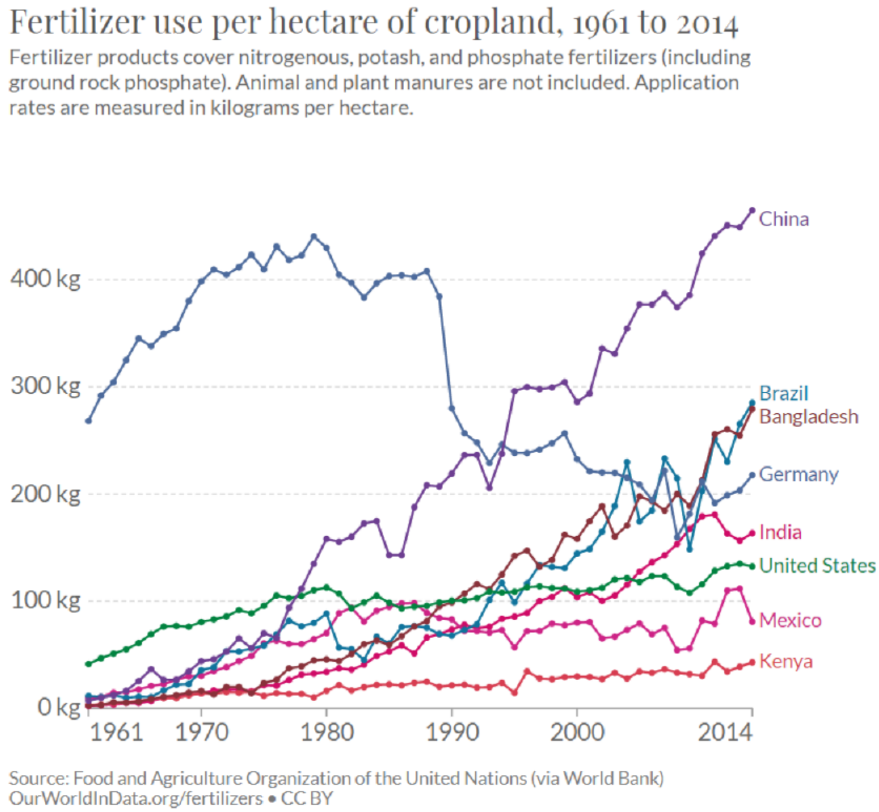
rus/potassium fertilizer usage, and increased yields and profits by roughly 5-7%. We find direct and indirect evidence on two essential mechanisms: 1) the salient feature of the technology leads to the overestimation of the return to that technology, resulting in the overuse of it, and 2) the overestimation of the return to the salient technology leads to undervaluation and underuse of the technologies with low salience. Our interventions significantly 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, adopt either the mobile application that provides customized recommendations or the leaf color chart intervention, then the total usage of nitrogen fertilizer would be reduced by 1.76 million tons (related to $N_2O$ emission reduction), $CO_2$-equivalent emissions could be reduced by 37.4 million tons, and yield would rise by 11 million tons (equivalent to 22 billion Chinese Yuan and 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 - 23.4 USD). In terms of the mobile application, the one-time cost for the programming is roughly 30,000 RMB and the annual cost of server maintenance is roughly 10,000 RMB. With regard to the leaf color charts, each only costs 8-30 RMB. If we apply the soil testing and mobile application intervention 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 land plots of 7.5 mu.

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 1.1, 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 causing nitrogen overuse. In the 1990s, Many farmers just followed the extension agents' advice. However, it's very likely that the extension agents recommended the application level under output maximization to follow the national food security strategy proposed by the central government. Under such food production target, smallholders might only respond to the average return of fertilizers, rather than the marginal return. 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 supported by 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 to help other developing countries avoid following the same fertilizer misuse path.

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.

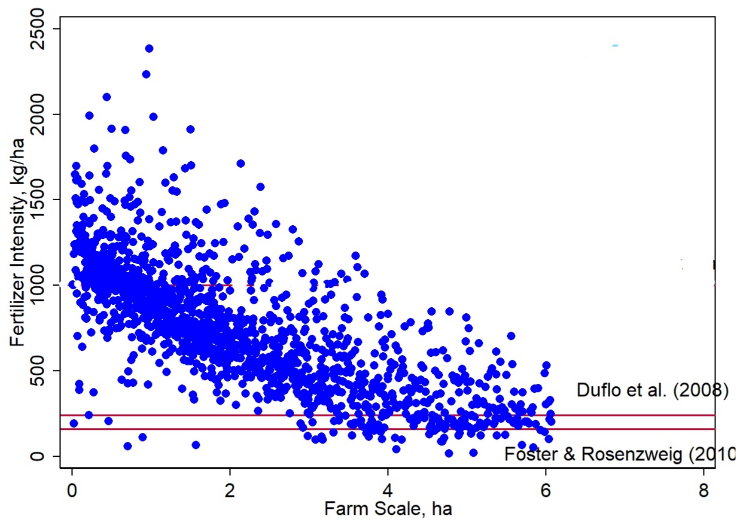
Figure 1.1: Cross-country Fertilizer Intensity, kg/ha



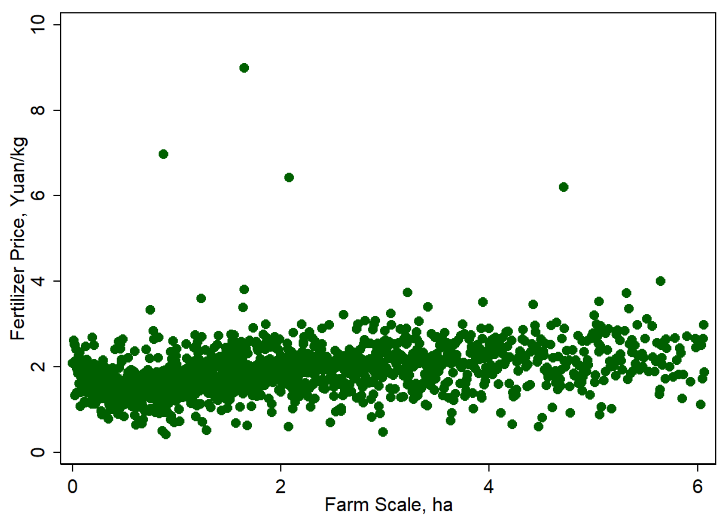

*Note:* This figure presents the pattern of cross-country fertilizer use per hectare from 1961 to 2014. The selected countries, in terms of fertilizer intensity in 2014, from top to bottom are China, Brazil, Bangladesh, Germany, India, United States, Mexico and Kenya. Most developing countries are on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

Figure 1.2: Fertilizer Intensity (kg/ha) and Price by Farm Scale

(a) Fertilizer Intensity (kg/ha) by Farm Scale



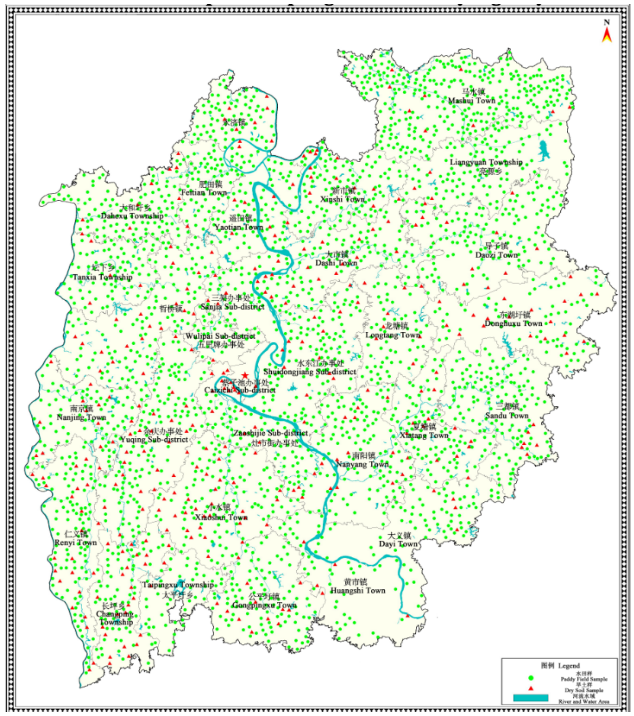

(b) Fertilizer Price by Farm Scale



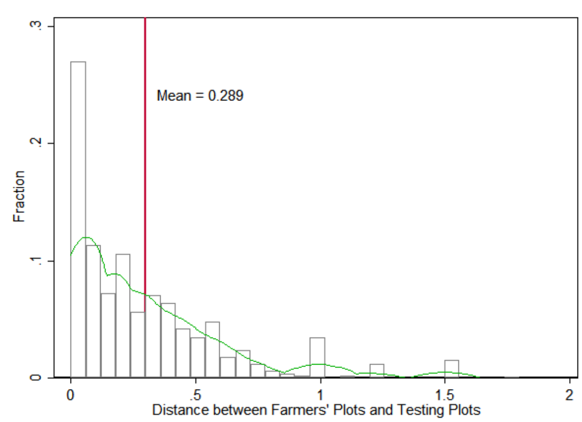

*Note:* The top panel of this figure depicts the relationship between fertilizer use per hectare (kg/ha, vertical axis) by farm-scale (horizontal axis) using a rich panel dataset which traces roughly 20,000 farming households from 1986 to 2015 in China. The top red line indicates the optimal level of fertilizer use in western Kenya in Duflo, Kremer, and Robinson (2008), roughly 242 kg/ha, and the bottom line shows the optimal level in India by Foster and Rosenzweig (2010).
The bottom panel shows the distribution of fertilizer price by farm scale.

Figure 1.3: Link Farmers to the Universal Soil Analysis Points

(a) Universal Soil Testing Program in Leiyang, Hunan

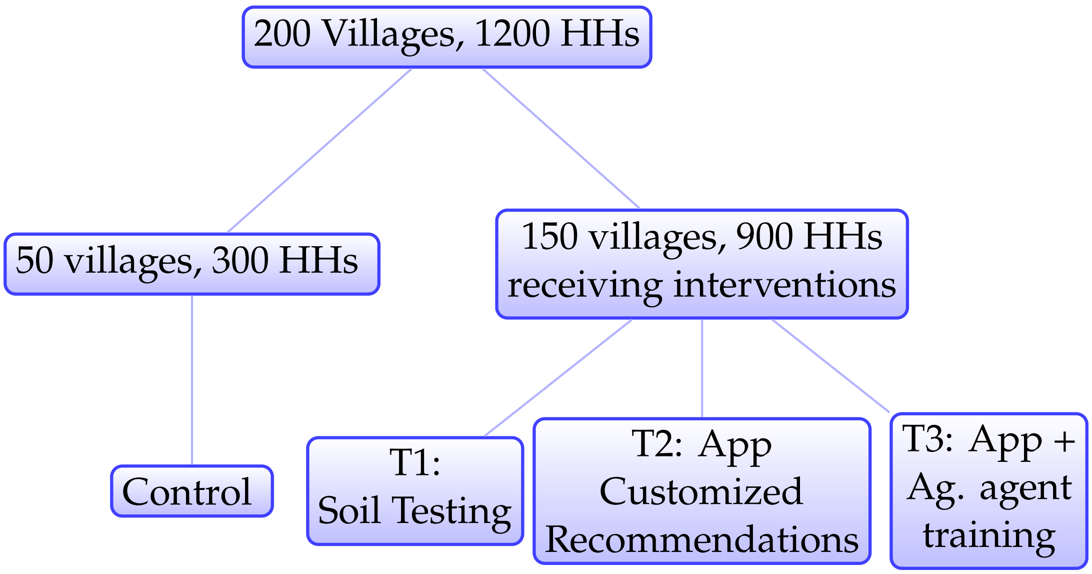


(b) Distance between Farmers' Plots and Nearest Testing Plots



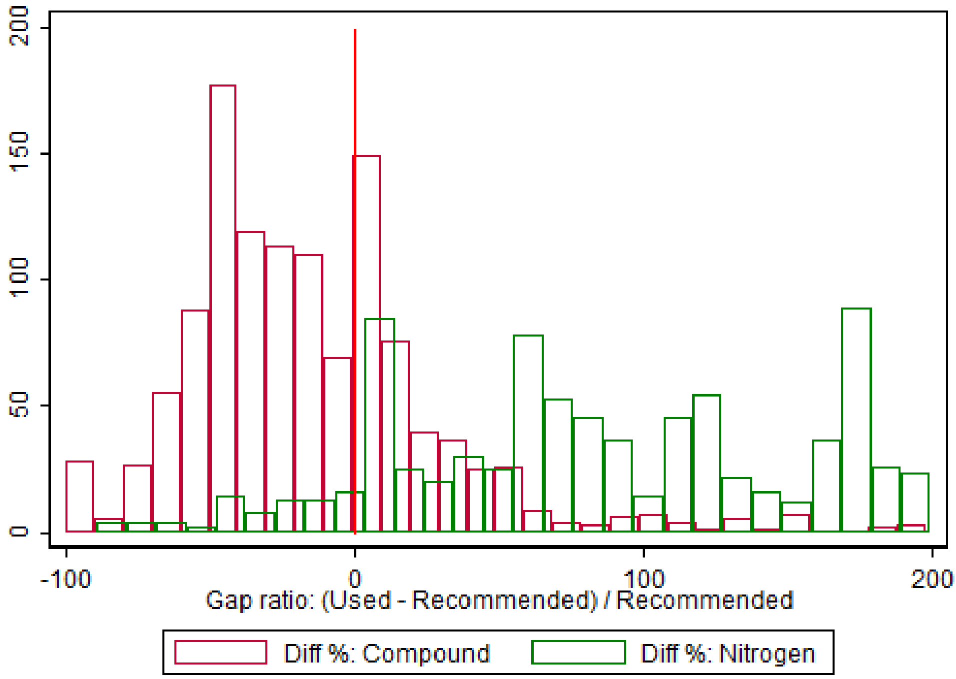

*Note:* This figure presents (a) the universal soil analysis in Leiyang, Hunan and (b) the distribution of distance between farmers' plots and the nearest testing point. In Figure 1.3a, each green dot is a testing point where a piece of soil sample was collected and analyzed in the lab for micronutrients 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 1.3b 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 1.4: 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.

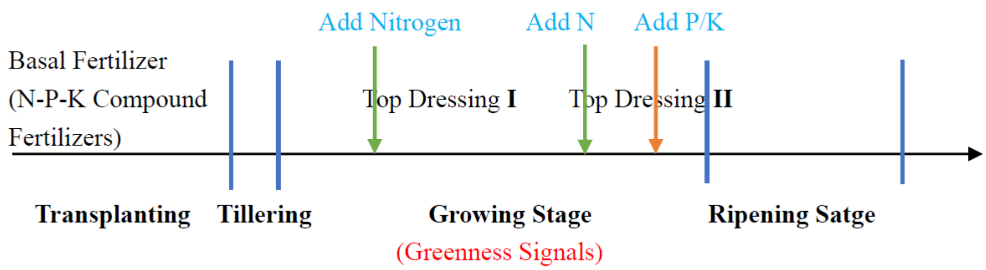
Figure 1.5: Gap in Fertilizer Application: (Used- Recommended)/Recommended



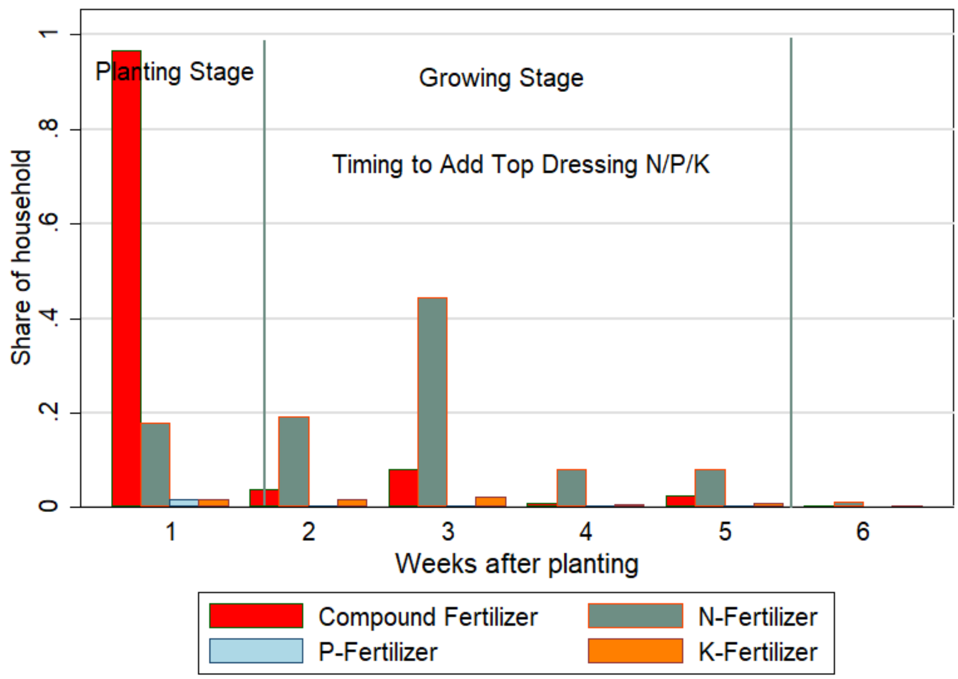

Note: This figure demonstrates the Gap in different fertilizer use between the actual use and the recommended use based on soil analysis data. The y-axis indexes the number of households, and the x-axis indexes the deviation percentage points, which equal to (actual Use - recommended)/recommended. The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

Figure 1.6: Actual Application Deviates the Recommended Practice in Growing Stage

(a) Recommended Practice by Time




(b) Actual Practice by Time



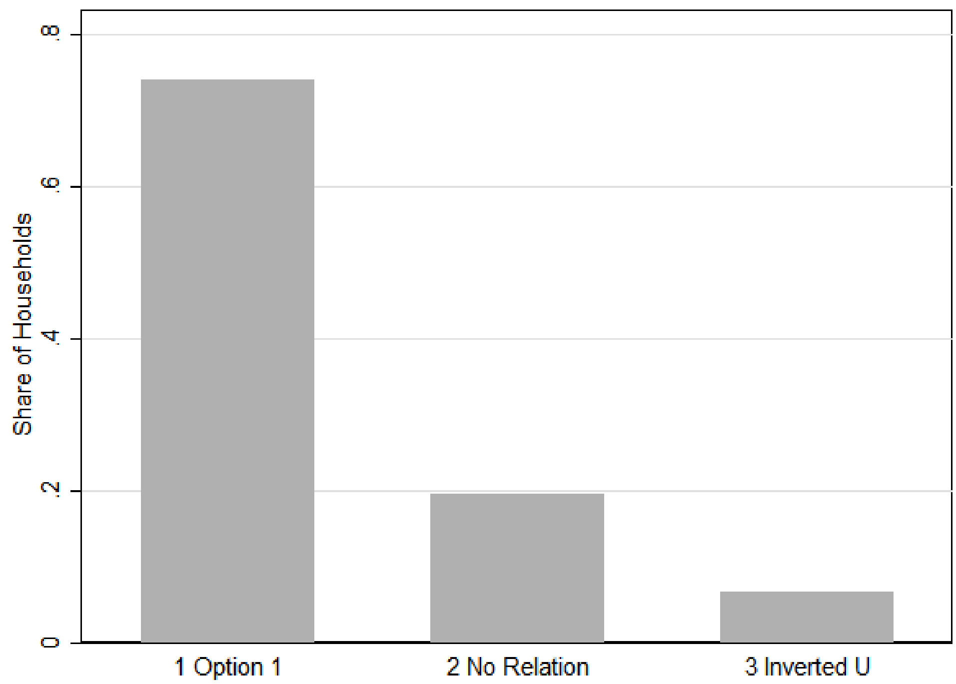

Note: These figures present how the actual fertilizer use from survey deviates the recommended practice. Figure 1.6a shows the recommended fertilizer application, divided into two main stages. During the transplanting stage, farmers are recommended to apply N-P-K compound fertilizer as the basal fertilizer, while during the growing stage, top-dressing nitrogen and phosphorus/potassium fertilizers are suggested to be used. In Figure 1.6b, the y-axis indexes the proportion of household applying different fertilizers, and the x-axis indexes different time. The left panel of Figure 1.6b is consistent with the recommended practice as shown in Figure 1.6a that almost all of the farmers were applying N-P-K compound fertilizer in the planting stage. However, the middle panel of Figure 1.6b shows that farmers only adjusted nitrogen use, but never added any phosphorus and potassium, which deviates the recommended practice.

Figure 1.7: Overestimation of the Return to Greenness

Option 1: The greener, the higher the yields are

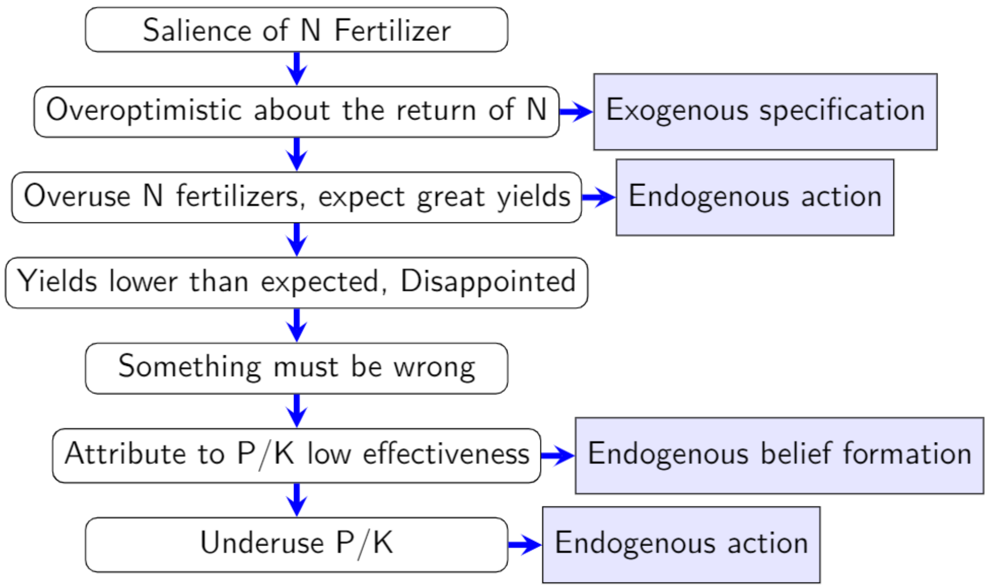
Option 2: No strong Relationship

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.

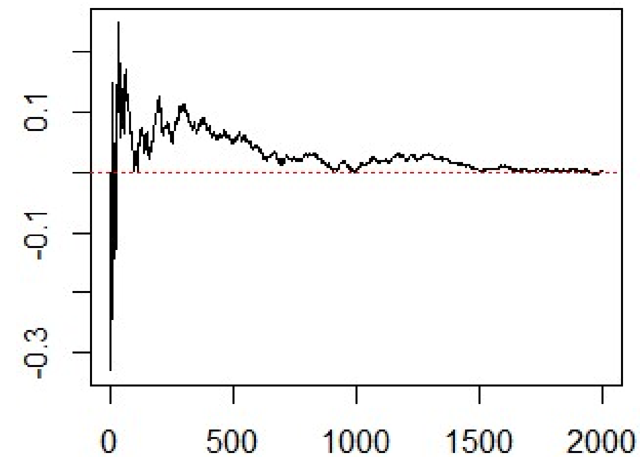
Figure 1.8: Model Intuition: Misspecified Learning Process



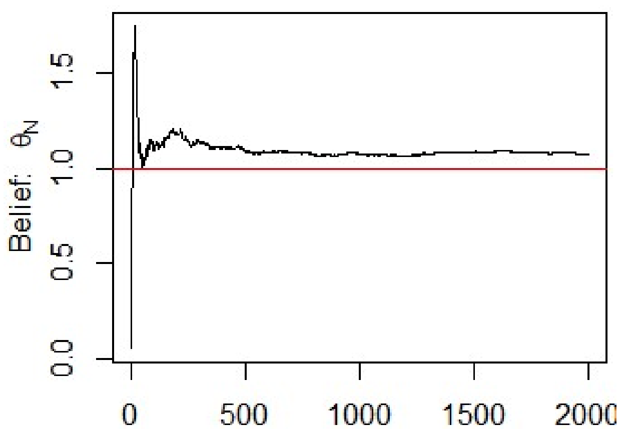

Note: This figure presents the basic logic of our model intuition.

Figure 1.9: Convergence of Beliefs about the Effectiveness of Different Fertilizers

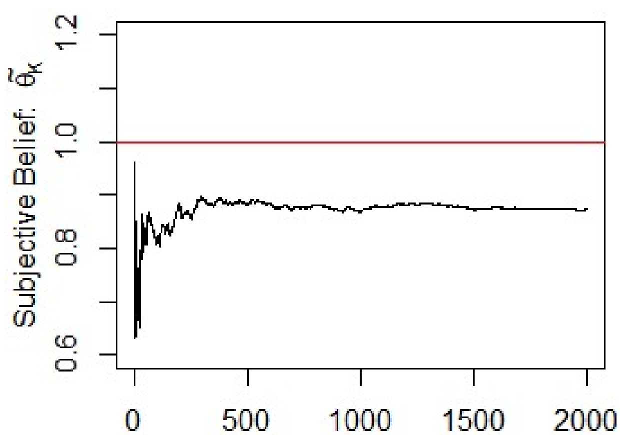
(a) Convergence of $g$ Function ($g$ = 0 ⇒Equilibrium)




(b) Learning Effectiveness of Nitrogen

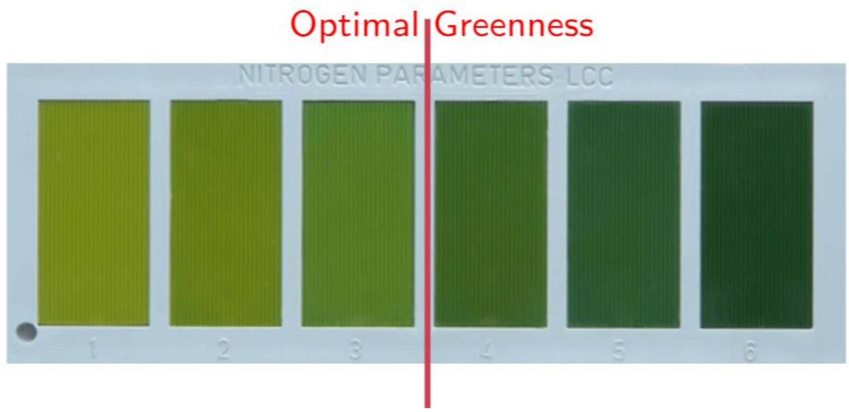



(c) Learning Effectiveness of Phosphorus/Potassium



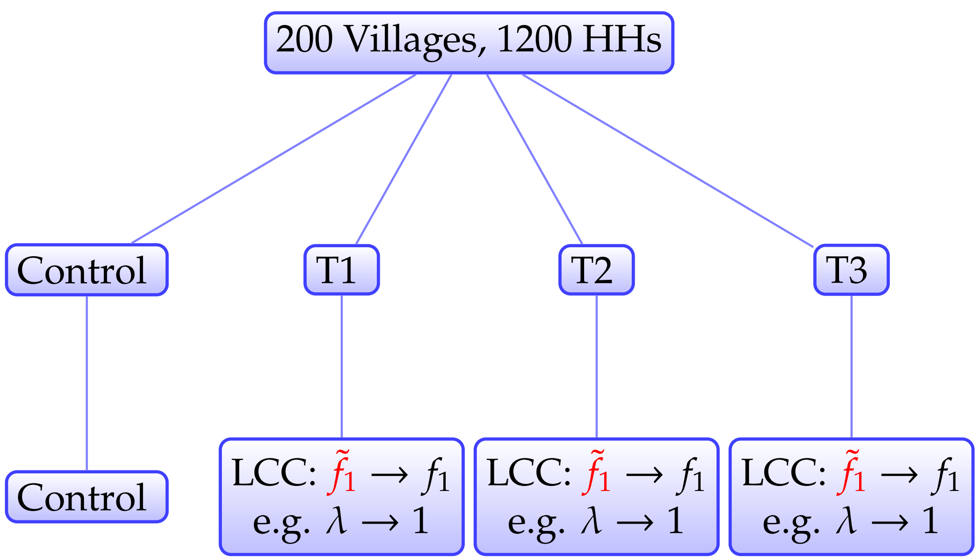

Note: This figure presents simulations about farmers' Bayesian learning on the effectiveness of nitrogen- v.s. phosphorus/potassium fertilizers. The top panel shows the convergence of $g$ function, indicating the new equilibrium. Figure 1.9b displays farmers' belief about the effectiveness of nitrogen fertilizer compared to a true value of 1. Figure 1.9c depicts farmers' belief about the effectiveness of phosphorus/potassium fertilizer relative to a true value of 1, which is slower and undervalued.

Figure 1.10: A Sample of Leaf Color Chart



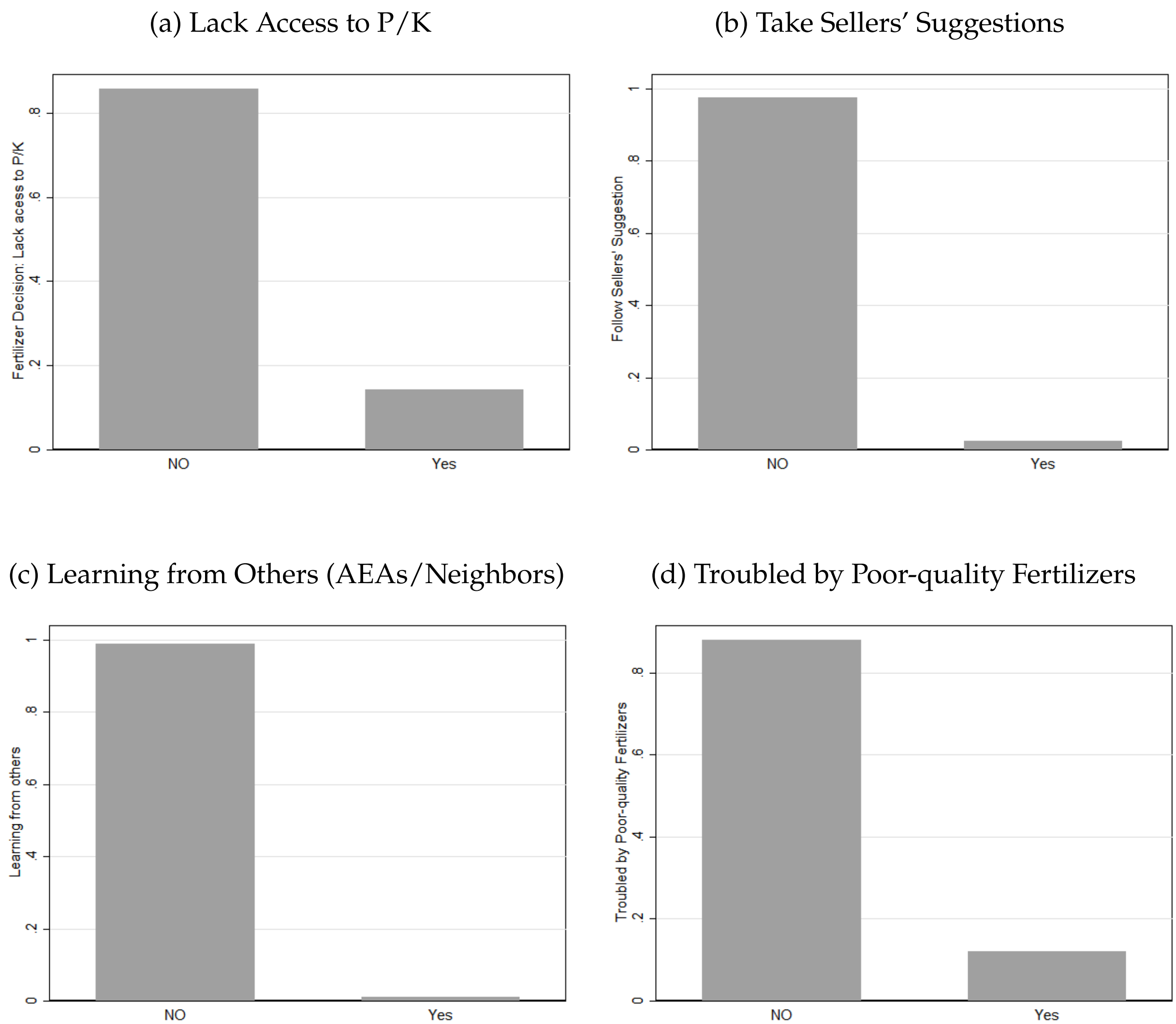

Note: The figure presents the six-color leaf color chart and the suggested optimal greenness.

Figure 1.11: 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 1.12: Supply Side-factors on Fertilizer Use




(a) Lack Access to P/K

(b) Take Sellers' Suggestions

(c) Learning from Others (AEAs/Neighbors)

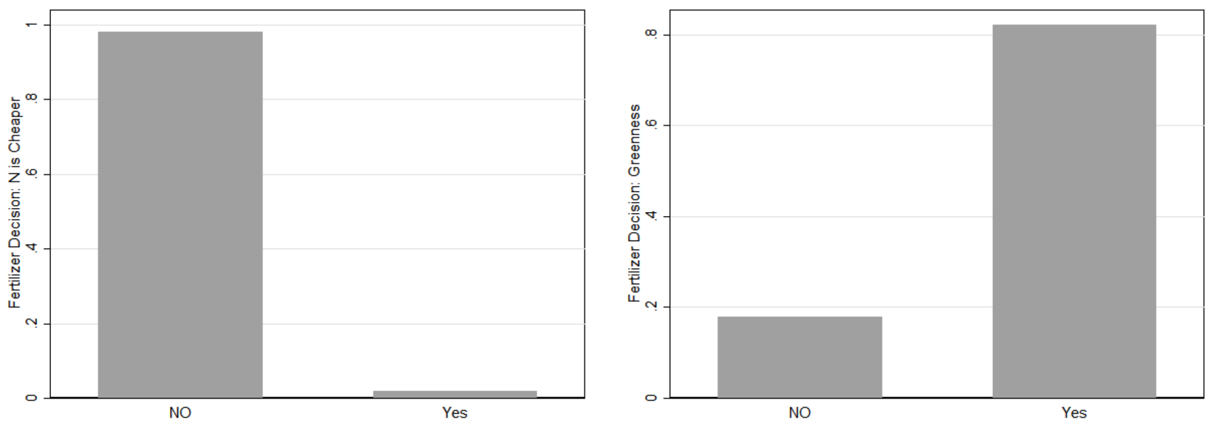
(d) Troubled by Poor-quality Fertilizers

*Note:* The figure presents supply-side alternative explanations that might affect fertilizer usage. Figure 1.12a shows that farmers were not subject to the supply constraints of phosphorus and potassium. Figure 1.12b rejects the hypothesis that fertilizer sellers influenced farmers' choice of different fertilizer use. Only lower than 3% farmers followed the recommendations from sellers. Figure 1.12c suggests no evidence on learning from others, while Figure 1.12d shows direct evidence that quality was not a concern for farmers' choice.

Figure 1.13: 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 1.13a). Figure 1.13b shows direct evidence on farmers' decision-making during the growing stage—based on greenness signals.

Table 1.1: 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 1.2: 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 |
| Observations | 1177 | 1177 | 1177 | 1177 | 1177 | 1177 | 1177 | 1177 | 1177 |

*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 1.3: 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 |
| Observations | 1177 | 1173 | 1177 | 1177 | 1177 | 1177 |

*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 1.4: 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 |
| Observations | 1177 | 1177 | 1177 | 1177 | 1177 |
| R squared | 0.111 | 0.00231 | 0.0999 | 0.134 | 0.172 |

*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 1.5: 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 1.6: 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 |
| Observations | 1153 | 1153 | 1153 | 1153 | 1153 | 1153 | 1153 | 1153 | 1153 |
| ***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 1.7: 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 |
| Observations | 1153 | 1153 | 1153 | 1153 | 1153 |
| R squared | 0.128 | 0.002 | 0.106 | 0.128 | 0.134 |

*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 1.8: 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/ | -4.661*** | 2.936*** | 2.688*** | -5.109*** | 3.070*** | 2.913*** |
| App + Training + LCC | (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 1.9: IV Estimation: Deviation in Fertilizer Application and Yields

| | (1)<br>IV-First Stage<br>$logGap^2$ | (2)<br>2SLS<br>Yields | (3)<br>2SLS<br>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 (1.6) and (1.7). 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 1.10: IV Estimation: Deviation in Fertilizer Application and Revenues/Costs

| | (1) IV-First Stage $logGap^2$ | (2) 2SLS Revenues | (3) 2SLS Fertilizer Cost | (4) 2SLS 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 (1.6) and (1.7). 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%.

# Chapter 2

# The Causal Impact of Economics Education on Decision-Making

## 2.1 Introduction

Does education impact decision-making? Evidence for the causal impact of education on decision-making is mixed. On the one hand, Banks, Carvalho, and Perez-Arce (2019) find that an additional year of required education has no significant impact on the quality of decision-making. On the other hand, a randomly assigned financial education program seems to successfully improve students' decision-making (Kim et al., 2018; Lührmann, Serra-Garcia, and Winter, 2018). Such mixed results raise the idea that the content of the education—in particular, studying economics—may determine whether education has a causal impact on decision-making.

We follow this line of inquiry by considering a setting in which students educated in the same college are quasi-randomly assigned to different majors. We hypothesize that decision-making skills change as the exposure to an economics curriculum increases, similar to the effect of curriculum in other contexts (Cantoni et al., 2017). While a strand of literature has attempted to distinguish between learning and the selection effects of economics education, primarily on social preferences (Marwell and Ames, 1981; Carter and Irons, 1991; Kagel, Kim, and Moser, 1996; Faravelli, 2007; Fisman, Kariv, and Markovits, 2009; Bauman and Rose, 2011), existing evidence on the presence of causal effects is mixed and could be improved in at least two aspects. First, even with longitudinal data, it is inherently difficult to rule out the selection problem induced by individuals' major choices. Second, while previous studies usually investigate a particular aspect of students' decision-making, an economics education could lead to fundamental changes in risk and several aspects of social preferences, which have not been thoroughly studied.

In this study, we overcome these limitations to evaluate the causal impact of economics education on decision-making by exploiting a unique institutional setting in China where admission is determined by students' scores on the College Entrance Exam (CEE). Due to the assignment rule of this centralized admission system,[1] the distribution of CEE scores among those who end up in the same major are highly concentrated. In the college that we investigated, for example, the standard deviation of scores among the students is less than .1 times the standard deviation of those who took the College Entrance Exam. Thus, those students who ended up studying economics/business within this educational system did so simply because their CEE scores were marginally higher than some competing students. Accordingly, we study the comparable sample created by this environment to evaluate the decision-making behaviors of those students who end up studying economics/business and those who take an alternative major because of marginally lower CEE scores. This key variation allows us to identify the causal effects of an economics education on the decision-making skills of comparable students close to the economics-admission cutoffs.

To capture these decision-making choices, we analyze the data of a survey conducted among students near the economics cutoff scores in a Chinese university, the Central University of Finance and Economics (henceforth CUFE), where a considerable number of students have economics/business majors.[2] The university administrators, who were interested in the impact of majors on students, distributed an online survey to some college students at CUFE via Student Central (an online official campus Learning Management System (LMS) that provides a resource for instructors and students to enrich the teaching and learning experience). To compare students who narrowly met the cutoff of economics/business majors with those who did not, students were invited based on whether their scores were close to the cutoff for whichever economics/business major they applied for.[3] Students who received the invitation used assigned accounts to log in and fill out the survey via computer. The survey asked for students' responses along several dimensions of decision-making—i.e., risk preferences, social preferences, and probabilistic beliefs. These questions were elicited in an incentivized manner.[4]

Our findings convey three main conclusions. First, students in economics/business majors exhibit a significant change in risk preferences. Specifically, students receiving an economics education become more risk neutral compared to those who have the same major preferences but end up in a non-economics/business major. As risk neutrality in small-stakes situations is viewed as an expected utility-maximizing behavior (Rabin, 2000), our findings imply that an economics curriculum may induce students to behave as a consistent expected utility maximizer in small-stake gambles. Second, students in economics/business majors show higher decision-making qualities in the investment game, where probabilistic reasoning is essential. This positive finding in probabilistic beliefs sheds light on the possibility of debiasing statistical reasoning. Third, while our results show that economics/business students' behaviors in small-stake social preference games do not change on the whole, these students' perceptions of others appear to shift significantly. Specifically, economics/business students are more inclined to believe that

[1] See Chen and Kesten (2017) for a detailed introduction to the Chinese college admission system.

[2] We use economics/business to signify majors/programs of study often housed within either economics or business schools.

[3] See Section 2.2 for details about survey distribution.

[4] Students received sign-up compensation and payoffs in each module, based on their responses.

others give less in the Dictator Game, are less engaged in reciprocity in the Trust Game, and share less in the Trust Game. This finding suggests that the economics/business curriculum plays a subtle role in shaping social preferences: while individuals' altruism remains unchanged, an economics education appears to affect how individuals interact with other people, especially in cases where their actions depend on perceptions of others.

The remainder of the paper is structured as follows. Section 2.2 elaborates on institutional details. Section 2.3 discusses our empirical strategy. Section 2.4 presents our main results for the impact of economics education on students' risk preferences, social preferences, and probabilistic beliefs. Section 2.5 conducts a comprehensive set of robustness checks. Section 2.6 offers the conclusions and the limitations of our study.

## 2.2 Background, Data, and Institutional Details

### Enrolling in Majors within Chinese Colleges

The admission process for colleges/universities is centralized in China. Before graduating from high schools, students must take the College Entrance Exam (China's National College Entrance Examination, also known as Gaokao, henceforth CEE). The exam is held once a year on June 7th and 8th, and all students must take it in their province of residence. The CEE includes three mandatory tests—in mathematics, verbal Chinese, and verbal English—and two optional tests in liberal arts and sciences. The maximum score of the CEE is usually 750 points.

After receiving their scores, students are required to declare their college preferences through the centralized system, along with their major preferences within each preferred university. They can rank several universities (the maximum number of which varies from province to province but is usually between four and ten) and subsequently usually rank six choices of major within each university, in order of preference. The deadline for applying to colleges is typically at the end of June, and the admission process follows a college-then-major design wherein universities first admit students based on applicants' scores and college preferences, regardless of their major preferences. Thereafter, the admission process resembles a serial dictatorship, where the priority scores are almost completely determined by students' scores in the CEE [5]. Please see Chen and Kesten (2017) for a thorough description of the mechanism.

Major assignments start after students have been admitted to colleges. Students first need to submit a rank-order list (ROL) that contains up to six majors before the application. The ROL ranks from the most preferred to least preferred major. The assignment rule largely follows the Deferred Acceptance Algorithm (Chen and Kesten, 2019). In other words, each major will first specify an admission cutoff score based on their capacity and student demand (i.e., based on the ROLs that have been submitted), and each student will

[5]Exceptions include minority ethnic groups, an award in the National Olympiad for Math/Physics/Chemistry/Biology/Informatics, an Athletes Award, or demonstrating excellence in some extra exam held by colleges to search for students with special talent.

then be assigned to a major that satisfies the following two conditions: (I) her score is not lower than the admission cutoff score of the major; (II) she ranked the major highest among those that satisfy condition (I).[6] The algorithm starts with students applying to their most preferred majors, and these majors accept students according to capacity and effective score. Majors keep students with the highest score and reject the rest, at which time students who have been rejected by their most preferred major apply to their second choices, again using their effective score. Students who are denied their most preferred major apply to their second most preferred and go through round two selection together. Again, the highest scorers pass through, the remainders are arranged according to third-tier preferences, and so on. The algorithm continues until every student has either been admitted or has exhausted their preference list.[7]

In equilibrium, there will be a sharp admission cutoff (lowest admission score) for each specific major that has imbalance between supply and demand. Given that students' scores are highly homogeneous above and below the admission cutoff scores, conditional on their preferences, this assignment is a quasi-experiment that randomly assigns students to different majors despite their similarity in major preferences and academic ability.

## Logistics of Survey Distribution

The rules governing major assignments within this system imply that a causal effect of education on decision-making skills may be sussed out by comparing students who have the same major preferences but end up in different majors due to small differences in their CEE scores. Consequently, the sampling strategy of our online survey closely follows this conceptual comparison. For this study, university administrators categorized majors at CUFE into four groups: economics/business, law/sociology, natural sciences, and humanities/language studies. Each major was classified according to its corresponding curriculum. For example, students from economics/business majors receive far more training in economics compared to students in other non-economics/business majors (see Table B.2 and Table B.3 for details). After students' majors were classified, the administrators started to invite and distribute surveys to students if students had an economics/business major as their most preferred major in their rank-order list and if their scores were not far from the economics/business admission cutoffs.

Upon agreeing to participate, each participant was assigned a personal account to sign into the survey website and was invited to start the survey with any internet-connected

[6] The major assignment is determined by an "effective score". The effective score is exactly the exam score for the vast majority of students, with a few exceptions that are unobserved to us. What we know from conversations with members of the admission committee is that the effective score may be less than the original priority score when students put an extremely competitive major at the bottom of her rank-order list, but such operation is rarely executed and has been gradually abandoned in recent years. Effective scores might be more than the original exam score for reasons including ethnicity, an award in the National Olympiad for Math/Physics/Chemistry/Biology/Informatics, an Athletes Award, or demonstrating excellence in some extra exam held by colleges to search for students with special talent.

[7] At this point, the college will try to accommodate students' preference by assigning them to majors that are not on their list but are as close as possible to their preferences.

computer. To provide participants with reliable access to the survey, the research team built the survey on a collaborative website with CUFE's official website domain.[8] The survey had several modules, including demographic information, risk preferences, social preferences, and probabilistic beliefs. During the survey, backward trace or modifying previously recorded answers was not allowed. In addition, the website system automatically cached the data after each module was finished. In total, the online survey took roughly 50 minutes to complete.

To stimulate participation, we incentivized all the questions we anticipated analyzing, and we awarded students an additional 20 Chinese Yuan (CNY) for participating in the background survey. As incentives varied according to students' responses, the range of payoffs spanned 20 CNY to 422 CNY, with a 50 CNY mean payoff, which is 4-8 times the price of a standard meal at school. The money was directly deposited into students' university accounts at the time they finished the survey. The substantial financial incentive ensured that students had enough impetus to participate in the survey and truthfully report their preferences and beliefs. Together with the school's administrative endorsement, these incentives boosted the response rate—roughly 72% of invited participants finished the survey.

### Curriculum at CUFE

Economics/business majors are usually taught as an eight-semester program in the vast majority of Chinese universities. There are around 10,000 students registered in 30 departments and research centers at CUFE, which consists of 80 majors. On average, students in economics and business majors take 47% of their courses in an economics curriculum[9] (roughly 18 courses and 54 credits), followed by general interest courses (27%, roughly 31 credits[10]) and other optional courses. Most economics courses closely follow translated versions of standard US textbooks. Tables B.2 and B.3 summarize how many students take their compulsory courses by the end of each semester if they stick to the curriculum requirements.[11] Among economics students, the most relevant courses—including microeconomics, macroeconomics, and finance—are taken by the end of students' third semester.

### Match Between Survey and Administrative Data

To evaluate what impact economics education had on students' decision-making skills, we linked our survey data to university administrative data to obtain students' rank-order lists for preferred majors, enrollment in majors, date of birth, province of origin, and score on the CEE. By matching students' survey responses to the administrative data (the success rate of matching surveys with administrative data was 100% since all the survey

[8] The website address is `prelab.cufe.edu.cn`.

[9] Statistics and probability courses fall under this curriculum.

[10] These courses include English, computer skills, and politics.

[11] Data were taken from the 2017 curriculum schedule for CUFE college students.

samples were drawn from administrative data), we identified some general patterns for applicants' major preferences in our sample.

First, economics/business majors were generally the most preferred majors for the vast majority of students, but many students could not enroll in their preferred economics/business majors due to excessive demand. Panel A in Table B.1 quantifies the distribution of applicants whose 1st, 2nd,..., 6th preferences fell into the categories for economics/business majors, natural sciences, law/sociology, humanities/language studies. Due to our sample's priority for students who listed an economics/business major as their first choice, in Panel A, all 989 selected students had an economics/business major as their most-preferred major preference. Among these 989 applicants, 798 and 687 applicants selected a different economics/business major as their second-ranked and third-ranked major preferences, respectively, whereas other applicants listed a science, law/sociology, or humanities/language major as their next-preferred choice. The last row of Panel A shows how many of our sample's 989 students were finally admitted to an economics/business major, namely 493. This split between those who were admitted to an economics/business major and those who were not provides the quasi-experiment for our analysis.

Second, Panel B of Table B.1 illustrates which ranked major (e.g., first choice, second choice) individuals were ultimately admitted into. This panel demonstrates that 753 out of 989 students were admitted within their top four preferred majors.

## 2.3 Empirical Strategy

### Sample Construction

The empirical strategy we employ is similar to Kirkeboen, Leuven, and Mogstad (2016). Conceptually, our sample consists of students who prefer economics/business majors to other majors but who may have ended up in different types of majors due to small differences between their CEE scores. To understand how we implement this strategy, we provide an example in Panel A of Table 2.1. In our example, two students, Qian and Wu, are both accepted CUFE students, and the college is considering which major to assign the students. We can infer from Panel A that both students prefer economics/business majors to others: Their top three preferred majors all fit under the category of economics/business majors, whereas their bottom three preferred majors are all non-economics/business related. As we have discussed in Section 2.2, despite the same major preferences, Qian will be assigned to her 2nd selected major, accounting, whereas Wu will be assigned to her 4th major, Chinese language/literature. Thus, these students end up in different types of majors merely because Wu's CEE score is slightly lower. The cutoff that contributes to such a difference is the minimum cutoffs of their top three majors. The counterfactual of an individual like Wu would be that she would have been assigned to an economics/business major as long as her CEE score met the cutoff of at least one major in her top three choices. Therefore, the effective cutoff in this case is 631.

To compare students accepted to their preferred major against those not accepted, we have to establish effective cutoffs relevant to their sorted placement. As we have discussed in Section 2.2, the vast majority of CUFE students prefer economics/business majors to other majors. In such cases, the effective cutoff is simply the minimum cutoffs of the economics/business majors that they put on the top of their ROLs. Additional complication arises when students do not always put one category of majors before others. In these cases, we instead consider the local ranking around the major that students are admitted to rather than the whole ROL (global rankings). To clarify what local ranking means, consider another two hypothetical students, Lin and Wang, whose ROLs are presented in Panel B of Table 2.1. Lin and Wang ranked economics/business majors 1st, 3rd, and 5th, and non-economics/business majors 2nd, 4th and 6th. Since Lin is admitted to law (2nd choice), the local ranking for him is finance (1st choice) > law (2nd choice). He would have entered his first ranked major, an economics/business major, if his CEE score met 635. In this case, therefore, the effective cutoff for Lin to enter an economics/business major is 635. However, in Wang's case, he is admitted to his 3rd choice, an economics/business major. The local ranking for him is economics (3rd choice, economics/business major) > marketing (4th choice, economics/business major) > journalism (5th choice, non-economics/business major). He would remain in an economics/business major as long as his score meets 626. In this case, therefore, the effective cutoff for Wang is 626.

To verify that the effective cutoff is indeed the key determinant in one's major, we plot the frequency and admission probability as a function of distance to the cutoff score. In Figure 2.1, we plot the probability of being admitted to an economics/business major against the distance to the cutoff score. Below the cutoff score, the probability of entering economics/business majors is 0, which confirms the admission rule that meeting the cutoff is a necessary condition for admission. The admission probability goes up from zero to nearly 1 after a score passes the admission cutoff. The few exceptions where non-admission occurs even though students' scores met the cutoff are perhaps caused by some special cases we cannot observe in our dataset.[12] Including these small fraction cases within our sample barely affects our results.

Our main analyses of the decision-making survey results focus on students whose CEE scores are close to the effective cutoff. Specifically, a student was included in our sample only if the distance between her CEE score and the effective cutoff was less than 0.1 times the standard deviation of the CEE score's distribution.[13] Within our analyzed sample, all the students preferred an economics/business major to other majors, at least locally. However, despite the same major preferences, some students ended up in an economics/business major simply because their CEE scores were slightly higher than

[12]We talked to administrators on the admission committee and learned there are several possible reasons that could contribute to non-admission. The most important reason relates to instances where students have the same score as the cutoff score but where capacity is constrained. The admission system will use students' score in a particular subject (usually math) to break ties and determine who will be admitted. See Appendix B3 for other possible reasons.

[13]We alter the selection criteria in Section 2.5 to test the robustness of our main results.

others. This arbitrary difference in educational treatment (e.g., economics education or non-economics education) serves as our key independent variable.

## Balance Test

Having constructed the target sample, we conducted some balance checks to examine whether there are any systematic correlations between major and other covariates. The results of the balance test are presented in Table 2.2 and show that predetermined covariates are quite balanced in the survey between the treatment group (students who ended up in an economics/business major) and the control group (students who ended up in other majors).[14] Table 2.2 shows that in the survey, 64% of non-economics participants are female, while this number is 62% in general economics fields.[15] The differences in column (4) in Table 2.2 as well as the standard errors in parentheses indicate that there is no substantial difference in these determinants between economics and non-economics students except for pre-college rankings. We find that pre-college rankings for economics students (top 11.86%), on average, is slightly higher than that of non-economics students (top 13.26%). This difference is perhaps driven by the construction of our regression sample: students in economics are on the right-hand side of the admission cutoffs while students in non-economics are on the left-hand side. As a result, economics students on average perform slightly better than non-economics students in high school.

## Specification

The aim of our empirical strategy is to estimate and interpret the coefficients of the following equation:

$$y_i = \alpha + \gamma econ_i + \sum_{k=2}^{k=6} \mu_{ki} + \epsilon_i \tag{2.1}$$

Where $y_i$ denotes decision-making choices for individual $i$. $\mu_{ki}$ is a vector of dummies that denotes whether students' $k^{th}$ majors in their rank-order list belong to the economics/business major category.[16] This set of dummies helps us control for the intensity of students' preferences regarding studying in economics/business majors despite the fact that, within this sample, all choose their most preferred major as an economics/business major. $econ_i$ indexes whether student $i$ is in an economics/business major or not based on our selected sample. $\gamma$, therefore, is our coefficient of interest, which measures the impact of the economics education on our outcomes of interest. We use this specification

[14] These alternative factors include gender distribution, father's education, mother's education, pre-college ranking, monthly consumption, monthly allowance from parents, and years of boarding before college.

[15] In the administrative data documenting all the students at CUFE, 62% are female.

[16] The first major application is always economics due to the design of our sample selection.

as a benchmark in our empirical analysis. In the following subsections, we analyze the outcomes of risk preferences, social preferences, and probabilistic beliefs, respectively.

## 2.4 Main Results

### Risk Preference

We focus on the two multiple price lists (MPL 1 and MPL 2, hereafter) in the risk-preference module, where students were asked to make choices between two options for a series of questions. In MPL 1, individuals were asked to choose between a series of monotonically increasing certain payoffs, {25 RMB, 30, 35,..., 55, 60} and a fixed lottery {30 with P=0.25 and 60 with P=0.75}. The place where students switch from the fixed lottery to a certain payoff indicates students' risk preferences. For example, students with a switching point equivalent to the certain payoff 35 are more risk averse than those equivalent to 50. To test whether the endowment effect is present—whether students value the lottery more when the question is framed as eliciting Willingness to Accept (WTA) the lottery (Table B.10 in Appendix B1)—we cross-randomized half the participants (480) into the WTA mode of questioning (Table B.10 in Appendix B1) and the other half (509) into Willingness to Pay (WTP) mode (Table B.11 in Appendix B1).

The second scenario used to elicit students' risk preference asked students to decide between two different lotteries. (Hereafter, we call this decision-making problem MPL 2 (Table B.12 in Appendix B1). In this scenario, Option A receives 30 RMB with probability $Pr = 0.25$, and 60 with $Pr = 0.75$. The series of Option B receives 400 with increasing probabilities $Pr(400) = \{0.01, 0.03, 0.05, 0.07, ..., 0.23, 0.25\}$, or receives nothing, with probability $1 - Pr(400)$. Compared to MPL 1, MPL 2 provides subjects with more opportunities to exhibit some risk-loving preferences, which are not encouraged in economics textbooks but quite common in gambling. For example, those who choose Option B given a probability of winning less than 0.13 are opting for the riskier option even when its expected payoff is also lower.

We start by analyzing MPL 1. The proportion of students exhibiting multiple switching in our sample is low (1% for non-economics students and 0.7% for economics students). In our analysis, we exclude students who give multiple switching responses and focus exclusively on students who give consistent answers, with at most one switching. Let $p$ denote the price that students are willing to forgo in exchange for the fixed lottery. The first two columns of Panel A in Table 2.3 summarize the interpretation of students' choices in MPL 1.

We first analyze the difference in the distribution of responses for MPL 1, pooling both WTA and WTP together in Figure 2.3 (the coding of the switching point is shown in the third column of Table 2.3), where the red line in each histogram indicates the switching point that indicates risk neutrality. We can see from the figure that economics students are more risk tolerant, and substantially more economics students are risk neutral.

Based on the interpretation of the switching point in the second column of Table 2.3, students are categorized into four groups: dominated choices, risk loving, risk neutral, and risk averse. We then conduct regression analyses to investigate the causal effect of economics education using the specification in equation (1) in Section 2.3. Columns (1)-(3) of Table 2.4 present the estimation results. Column (1) estimates the impact of an economics/business major ($Econ = 1$) on the proportion of students who appear to be risk neutral. Column (2) measures the impact of an economics/business major on the share of risk-loving students. Column (3) compares the impact of an economics/business major on the proportion of risk-averse students. In these regressions, we exclude people who make dominated or inconsistent choices, and we limit to students who put both an economics and a non-economics/business major in their rank-order list (henceforth, common support of major preference[17]) but who preferred economics/business majors more. Taken together, the results suggest that roughly 11.8% of students who would be risk averse without economics education become risk neutral when making choices in MPL 1.

We additionally examine whether the Willingness to Accept/Willingness to Pay (WTA/WTP) framing affects students' risk preferences. Table 2.5 presents the main analysis using separated sub-samples. The first, third, and fifth columns include students who answer MPL 1 under the WTA mode, whereas the second, fourth, and sixth columns include students who answer MPL 1 under the WTP mode. Under the WTP mode, the proportion of risk-averse and risk-neutral economics students does not significantly differ from non-economics students, suggesting that the economics education does not have a significant impact on students' decision-making when loss aversion is at play. This finding could provide support for the notion that loss aversion is inherent and cannot be eliminated (Chen, Lakshminarayanan, and Santos, 2006). In contrast, under the WTA mode, economics students value the lottery more compared to non-economics students, suggesting that the effect of cation is most salient when the framing effect that stems from loss aversion is muted.

Next we turn to MPL 2 and relate the results to what we have found in MPL 1. The commonality between MPL 1 and MPL 2 is that both lists have two options for students to choose, and Option A is the same fixed lottery (win 30 w.p. 25%, win 60 w.p. 75%). For the purpose of exposition, we denote the winning probability required for students to choose the lottery with payoff of 400 Yuan by $p$. The first column in Panel B of Table 2.3 describes the choice in MPL 2 and the second column reports the interpretation of the choice. The third column in Panel B of Table 2.3 presents the coding of the choice.

We start with plotting the empirical distribution of the choice for MPL 2 in Figure 2.4, where the vertical red line again indicates the risk-neutral choice. The higher the switching point is, the more risk averse subjects are. There is a more significant spike on risk-neutral choices, particularly among economics students. The primary change in the distribution of the risk preferences manifests among students who used to be risk-loving (i.e., switching point $< 15$) but who become risk neutral, which we confirm in Columns (4)-(6) of Table 2.4.

[17]Students who had at least a non-economics/non-business major as their choice.

Similar to the findings in MPL 1, in MPL 2, we again find that the proportion of risk-neutral economics students is significantly higher than risk-neutral non-economics students, with the proportion of risk-loving being lower among economics students. Combined, these results suggest that an economics education may decrease students' risk-taking behaviors in an environment where taking more risks would lead to less expected values.

Taken together, both MPL 1 and MPL 2 suggest that economics education induces students to behave as risk-neutral agents. As Rabin (2000) argues, relative to total wealth in a whole life cycle, a modest reward such as what we offered in our study should be viewed as small-stakes gambles. Our preferred interpretation of the finding is that an economics education induces more students to behave like consistent expected-utility maximizers, who arguably have a higher quality of decision-making. There are, of course, other interpretations that are hard to rule out with the data we have. For example, if students believe that there exists "a right answer" and take the survey in "exam mode," those with an economics background might be better at finding the supposed right answer because of their training. Alternatively, students with an economics background are conceivably more likely to try harder when answering such familiar decisions questions thanks to their courses. While we do not take a stand on which interpretation is correct in our case, there is no doubt that students who receive an economics education are more likely to behave as if they are making "consistent" choices, and such changes could possibly affect how decisions are made in other situations.

### Social Preference

In the module of social preferences, students were asked to play a series of real-stakes games, wherein they received the payoff promised if their responses were randomly selected for reward.[18] To measure students' social preferences, three social preference games were conducted in the survey: the Dictator Game, the Ultimatum Game, and the Trust Game. For each game, each student was randomly assigned one of three roles: Player A, Player B, or Bystander. In our context, Players A and B play the games and the Bystander answers questions about her beliefs regarding Player A's and/or Player B's actions.

In the Dictator game, Player A corresponds to the Dictator. We asked Player A the following question: How much money out of 500 Yuan are you willing to share with Player B (the "Receiver")? Player B corresponds to the Receiver, and students who were assigned to the role of Player B were informed of the game but did not need to take any action. We asked the Bystander the following question: As a bystander, what do you think is the median value of the Dictator's sharing value in the Dictator Game? In terms of monetary incentive, Player A will get 500 minus the amount she/he sends out, and Player B will get the money that Player A is willing to transfer. For the Bystander, the payoffs depend on the accuracy of her belief. Following the binarized scoring rule (Hossain and Okui, 2013)

[18]The monetary award in this module was designed to be particularly large to boost the survey's response rate. We made it clear in the survey that 20 participants' responses would be randomly selected for award.

as well as previous literature (Krupka and Weber, 2013), the rule is that the closer the belief is to the truth, the more likely the Bystander will be able to win a 500-Yuan award. The probability she wins the award is max{0, 1-(difference between belief and truth/150)}.

The results from this game are summarized in Table 2.6. On average, Dictators share about 190 Yuan out of 500 Yuan (38%) with the other player, and the amount of sharing does not differ significantly across economics and non-economics students, regardless of the regression specifications (Column (1) and (2)). Interestingly, when comparing Columns (1) and (2) (the Dictator's actual sharing) to Columns (3) and (4) (the Bystander's beliefs about the Dictator's sharing), we find that non-economics students' prediction (which is shown at the bottom of Columns (3) and (4)) pretty much aligns with our data on the Dictator's actual sharing (as shown at the bottom of Columns (1) and (2)). Economics students' beliefs about the Dictator's sharing decrease substantially relative to non-economics students, which leads to more inaccuracy and pessimism in beliefs about the Dictator's behavior. We interpret our findings as evidence suggesting that an economics education leads students to believe other people will behave in line with standard game theory predictions but economics education does not change what the student's own social preferences (e.g., altruism, social norms) may be.

In the Ultimatum Bargaining Game, Player A, the Proposer, received 500 Yuan, which she was told to split between herself and Player B in the first step. She could choose any amount to keep (from 0 to 500 Yuan), giving the rest to Player B, the Receiver. In the second step, Player B could choose to accept, which resulted in the same outcome as the Dictator Game, or choose to decline, in which case both players got 0. Player A was asked to propose the amount she would give to Player B, and Player B was asked the minimum payoff he would receive to not decline the proposal. The Bystander was asked to predict the median of the distribution of Player B's rejection threshold. Similar to the Dictator Game, Player A and Player B played the game and received the exact amount of money if their responses were selected. The Bystander was rewarded for the accuracy of her beliefs, with our payout rule saying that the closer the belief was to the truth, the more likely the Bystander would be able to win a 500-Yuan award. The probability she won the award was max{0, 1-(difference between belief and truth/150)}.

The results from this game are summarized in Table 2.7. On average, students' belief about the rejection threshold (Column (4)) is higher than the average of the actual threshold (Column (2)), and the Proposer is willing to share much more (Column (6)) compared to expectations (Column (4)). This result suggests that some Proposers may be extremely averse to being rejected for instrumental or psychological reasons. We do not find significant differences in the rejection threshold, the Bystander's beliefs, and the Proposer's sharing in this game when comparing economics students' behavior to their counterparts.

The Trust Game was approached as follows: Player A, the Sender, could choose to send $X$ amount of 500 Yuan[19] to Player B, the Receiver. She was also informed that what

[19] We intended to say $X$ out of 200 Yuan, but the typo was not identified until the survey had been issued. Consequently, in the reimbursement stage, we paid the selected students the amount stated in the survey.

she sent would be tripled when Player B received the money. Therefore, when Player A shared a value $X$ with Player B, our game would give $3X$ to Player B. Upon receipt of the money from Player A, Player B could choose to send $Y$ amount of $3X$ back to Player A. The Bystander in this game was asked about his beliefs about Player A's and Player B's behaviors, similar to the Dictator Game and the Ultimatum Game. Each student was asked three questions for this game, as detailed below.

In Question 1, students were randomly assigned to play this game as either Player A or the Bystander. Player A was asked to choose the amount of money to send, and the amount $X$ could be 50, 100, or 150. Bystander was asked to predict the mean of the distribution of $X$.

In Question 2, every student was asked about Player B's choices, namely, how much money Player B would like to give back to Player A upon receiving money from Player A, whose amount was $M$ hypothetically. $M$ could be 50, 100, or 150. For each student, each value of $M$ appeared with equal probability.

In Question 3, every student was asked to predict the average response in Question 2. Specifically, every student needed to predict the median distribution of the amount of money given back if Player A hypothetically gives Player B $M$ Yuan—and $M$ could be 50, 100, or 150. For each student, each value of $M$ appeared with equal probability.

The results of the Trust Game are summarized in Table 2.8. Columns (1) and (2) summarize our findings from Question 1 for the role of Player A and the Bystander, respectively. Columns (3) and (4) present regression analysis for Question 2 and 3, respectively, where we run the following regression:

$$y_i = \beta_0 + \beta_1 econ_i + \beta_2 (M - 100)_i + \beta_3 econ_i * (M - 100)_i + \sum_{k=2}^{k=6} \mu_{ki} + \epsilon'_i \tag{2.2}$$

Where $M$ is defined as the hypothetical amount of money Player A would like to share. We normalize the amount $M$ by subtracting 100 from the amount, and we denote this normalized amount by $M'$ (hereafter $M'$ indexes $M - 100$). The rationale behind this normalization is that in the survey, Player A has only three options for disbursement: 50, 100 or 150 Yuan. We interpret an amount exceeding 100 as a generous action and an amount falling short of 100 as an uncharitable action.

Column (1) in Table 2.8 analyzes how an economics education affects students' sharing behavior as the Proposer in the Trust Game. It reflects the Proposer's beliefs about the other player's trustworthiness and inclination to reciprocate when the Proposer shares more in the first stage. We find that an economics education significantly reduces the amount of sharing, and the magnitude is about 10% of the average sharing among non-economics students. These results are consistent with the finding in Column (4): While there is no significant difference between economics and non-economics students as Player B, when asked how much will be given back if Player A gives 100 Yuan—the middle (and perhaps neutral) action—economics students on average believe that generous sharing by the Proposer has less impact on the other players' reciprocated behavior. To recapitulate, economics students as Player A, on average, share less because they believe that Player

B is less likely to return their generosity. Column (3), on the other hand, informs us that there is actually no significant difference in terms of reciprocity behavior across economics and non-economics students. These three columns together are in line with the Dictator Game findings, where economics students do not change their own social preferences but rather change their views about how fellow students will interact with people. The insignificant result in Column (2) could be interpreted as a second-order belief about subjects' tendency to reciprocate, because the finding reflects beliefs about students' willingness to share, which, in itself, is reflective of the first-order beliefs about Player B's willingness to reciprocate.

Overall, while we do not find strong evidence that economics students change their social preferences (altruism, pro-sociality, norms, etc.), they do change their first-order beliefs about others' social preferences. This observation has significant implications for how economics education could potentially shape human interactions: While economics education may not have a strong effect on people's own social preferences, it may well change how people interact with others in many social activities, as learning economics induces students to regard others as homo-economicus. Due to the limits of our design and sample size, however, the estimates here are not as precise as the results in the risk-preferences module. We hope that future research could pursue this line of inquiry and investigate whether this finding extends to other contexts and larger samples.

## Outcomes: Probabilistic Beliefs

The survey also contained three questions on probabilistic beliefs. These three questions were similar in that they asked students to allocate their resources (30 virtual coins for all questions) between two Arrow-Debreu assets, A and B. If event A/B was realized, for each coin allocated to A/B, students would gain a lottery ticket that would yield 200 Yuan with probability 1%.

The first question tested knowledge of the law of large numbers: When flipping a fair coin 1,000 times, event A specifies that the coin's head would appear at least 530 times, and event B complemented event A (less than 530 times). Regardless of preferences, students should always allocate as many of their assets to event B as possible, as the law of large numbers indicates that the probability of event B is almost 1. The second question was a placebo test where event A and B happen with the same probability, and to the best of our knowledge, no psychological heuristics can be related to this question. The third question tested the representativeness heuristic: Flipping a fair coin 100 times, event A specifies that exactly 50 previous flips were heads, and event B is complementary to event A. While the probability of B is almost 1, students who are influenced by "exact representativeness" (Camerer, 1987) may overestimate the probability of A and allocate too many of their assets to event A. More details about these questions are presented in the section on eliciting probabilistic beliefs in Appendix B2.

We can see from Table 2.9 that the results are highly consistent with our hypothesis that economics students are more rational because they allocate significantly more coins to event B for question 1 and 3 (Column (1) and Column (3) in Table 2.9). In contrast, when there is no single optimal strategy, their behavior does not differ much from non-economics students in question 2 (see Column (2) in Table 2.9).

Overall, our results regarding the probabilistic beliefs module suggest that the statistical courses in economics/business majors endow students with the ability to understand and calculate probability. Still, it is unclear whether students can and are willing to apply this acquired skill to their real-life decisions, as students may be responding to the questions in "exam mode." Further study is required.

## 2.5 Robustness Check

In this section, we conduct three robustness checks to test the validity of our conclusions: (1) discuss heterogeneity in treatment effects and exposure to economic courses; (2) compare non-causal and causal estimates using extra survey samples; (3) test robustness to sample selection criteria. Additional results on all other robustness checks, such as control of financial status, disappointing effects, encoding of major preferences, and gender heterogeneity can be found in Appendix B3.

### Exposure and Treatment Heterogeneity

While we cannot tease out the effect of taking a specific course in the economics curriculum, as most students take multiple compulsory courses at the same time, we have sufficient variation in curricula to test changes in decision-making over time because most students take compulsory courses during freshman and sophomore year. Due to the limited power in our design, we will only be focusing on the heterogeneity effect on risk and probabilistic beliefs. Note that the social-preferences module's between-subject design significantly weakens the power, such that there are usually fewer than 400 students in each regression, making it infeasible to divide the sample again by stage-of-education. Therefore, we do not pursue heterogeneity analysis in social preferences for this subsection.

Since our survey was conducted in December, sophomore, junior, and senior students had already finished the first three semesters of their compulsory courses. Among economics students, most relevant courses—including microeconomics, macroeconomics, and finance—are taken by the end of the third semester, as shown in Table B.2 in Appendix B1. Thus, we divided economics students into two groups: freshman vs. post-freshman, and we ran the following regression to examine the heterogeneity in treatment effects among economics students:

$$Y_i = \kappa + \beta_1 econ_i * freshman_i + \beta_2 econ_i * post\_freshman_i + \theta X_i + \epsilon_i \qquad (2.3)$$

Where $Y$ is the subjects' response. The constant $\kappa$ is the average response for non-economics students. $post\_freshman$ indicates whether a student is a freshman or not, where $post\_freshman = 0$ if student $i$ is a freshmen. $\beta_1$ and $\beta_2$ measure the effects of economic education before and after the main economics courses, respectively. $X$ denotes the control variables, such as major-preference fixed effects under common support of major preferences. If our hypothesis is correct, we would expect that in the presence of significant treatment effects, $\beta_2$ would be larger in magnitude (hence, more significant) compared to $\beta_1$. Among the regressions with no significant treatment effects, we would expect both $\beta_1$ and $\beta_2$ to be insignificant.

The results are summarized in Table 2.10. The outcome variable in Columns (1) and (2) is the share of risk-neutral students in MPL 1 and MPL 2. Columns (3), (4), and (5) pertain to the probabilistic belief questions on the law of large numbers (LLN), two identical choices, and the probabilistic belief questions on Exact Representativeness (ER), respectively. Consistent with previous results and our hypothesis, in Table 2.10, $\beta_1$s, the education effects on freshmen between economics and non-economics students are barely significant across all the outcomes, as the freshmen have only taken three months of classes. However, $\beta_2$ is statistically significant in risk preference (MPL 1 in Column (1) and MPL 2 in (2)), the probabilistic belief questions on the law of large number (Column (3)), the probabilistic belief questions on exact representativeness (Column (5)), and the Bystander's belief in the Dictator Game (Column (7)). Columns (4) and (6) indicate that economics students show no difference relative to non-economics students in the indifferent-choice question (the second question in Section 2.4) and social preference.

## Causal vs. Non-causal Estimates

Here, we consider whether causal estimates differ substantially from non-causal estimates and to what extent the magnitude of learning effects compare to those of sorting effects. The construction of our sample limits our ability to assess the full extent of sorting effects, as most students who were invited to participate in the survey were those who chose economics/business majors as their most-preferred majors. To approximate non-causal estimates with existing data, we conducted two types of exercises, both of which are reported in Table 2.11.

In the first exercise, we re-ran the causal specification as reported in Section 2.4, but made two important changes: (1) The university also distributed surveys to those who were not admitted through the college entrance exam (CEE). Among the 1634 students who took the survey, there were 510 who were admitted through non-CEE channels.[20] In Section 2.4, these students were excluded in our causal specification but are included here to form a non-causal sample. (2) We excluded controls on students' major-preference fixed effects in the non-causal specifications here, as these characteristics capture students' preferences toward economics/business majors, a key factor that determines the process of sorting.

In Table 2.11, Panel A reports the results of our non-causal estimates, and we also replicate the corresponding causal estimates from Section 2.4 in Panel C for the convenience of comparison. Different columns correspond to different outcome variables, as indicated at the top of the table. We can see from the comparison that non-causal estimates differ

[20]These channels include, for example, a special college enrollment plan for rural/poor students.

from casual ones, though to a various degree across different outcome variables. The estimates for risk preferences become less statistically significant (Column (1) and (2)), suggesting that sorting effects are present in our context. Similarly, the effect for the ER heuristic becomes smaller and less statistically significant in the causal specification (Column (5)), which is consistent with the fact that non-causal effects are strengthened by both learning effects and sorting effects. In the social-preferences module, the non-causal estimates differ even more from causal estimates. Specifically, we find in non-causal estimates that an economics education is significantly associated with lower sharing (Column (6) of Panel A) but not with Bystander's beliefs (Column (7) of Panel A), which is exactly the opposite of what we find in causal estimates. The statistical significance in our non-causal specification suggests both that some existing findings on economists' selfishness could be explained by sorting effects of major application/assignment and that what economics education really shapes is students' perception about other people's social preferences.

In the second exercise of this section, we considered the following regression to double check the non-causal effects:

$$y_i = \alpha'' + \beta'' * \#PreferredEconMajors_i + \epsilon''_i \tag{2.4}$$

Where variable # *Preferred Econ Majors* represents how many economics/business majors a student $i$ put in her ROL. Since the variable # *Preferred Econ Majors* is a measure for the intensity of students' preferences towards economics/business majors, the coefficient of interest in this specification is $\beta''$, because it captures whether students' preference—the key factor that dictates the sorting process—correlates with the outcome variables of interest.

We report the regression results for all major outcome variables in Panel B of Table 2.11. In the risk-preferences module (Columns (1) and (2) of Panel B), stated preferences are significantly and positively associated with the outcome variables, and the sign of the estimates is consistent with the causal effects in Panel C. Similar findings emerge in Columns (3), (4), and (5) for probabilistic beliefs. For the social-preferences module, stated preference is negatively correlated with Dictator's sharing (Column (6) of Panel B), which suggests the existence of substantial sorting in this particular dimension and explains why the estimate is significant in non-causal specification but not in the causal estimations in Section 2.4. In sum, these results suggest that the intensity of preferences for economics/business majors is indeed strongly correlated with our outcome variables of interest and could potentially contribute to sorting effects. Therefore, sorting effects should be carefully controlled for if researchers would like to obtain a causal estimate of the effects of economics education on certain outcome variables.

## Criteria of Sample Selection

In Section 2.4, we restrict the regression sample to 0.1 standard deviations within the distribution of CEE to make the students in the treatment and control group more homogeneous and thereby more comparable. In this subsection, we conducted two additional empirical exercises to check whether our results is sensitive to sample selection.

The first exercise was to re-run the regression analysis using an alternative sampling criteria. Specifically, in this robustness test, we limited the sample to students lying in the 0.15 times the standard deviation within the distribution of the CEE score. By applying this new criterion, we obtained a new sample with 963 students, of which 495 were in economics/business majors and 468 were in non-economics/business majors.[21] We then tested the robustness by regressions using equations (1) and (2). The results are shown in Table B.8. We find that the magnitude and significance of most coefficients on economics/business major are robust after the inclusion of the additional sample. Perhaps due to our small sample size, the estimate is somewhat more sensitive to such inclusions in Column (7), the Bystander's belief about reciprocity in the Trust Games.

The second exercise examined RD-type figures that plot outcome variables against the difference between individual's CEE score and the cutoff score for economic majors. Compared to the first exercise, the advantage of this second exercise is that it does not rely on any assumptions on the level of treatment effects as a function of distance-to-cutoff scores. Therefore, the flexibility could unmask the potential heterogeneity in treatment effects and shed light on the sensitivity of regression estimates to bandwidth selection.

We focus on the case of risk preferences, social preferences, and probabilistic beliefs, and our main results are presented in Figures 2.5, 2.6, and 2.7. In Figure 2.5, we plot the share of risk-neutral students against distance-to-cutoff for MPL 1 and 2, respectively, and find that the discontinuity is substantial for both cases around the cutoff score. In Figure 2.6, we plot the outcome variables from the social-preferences module. For the Dictator Game, we can see from the first two plots of the top panel of Figure 2.6 that there is no visible "notch" for Dictator's sharing, but the discontinuity is present for Bystander's beliefs about others' sharing for the Dictator game. These findings are consistent with our causal estimation results from Table 2.6.

In the two graphs of the Trust Game (the third plot in the middle panel and the plot in the bottom panel of Figure 2.6), we find that there is discontinuity for the Proposer's sharing behavior but not for the Bystander's beliefs about sharing. Findings in both the Dictator and Trust Games are in line with our previous regression conclusions from Table 2.6 and Table 2.8.[22] For the Ultimatum Game (the third plot in the top panel and the first two plots in middle panel of Figure 2.6), consistent with our regression results, the discontinuity is absent for Bystander's belief and Proposer's sharing. The discontinuity seems to be significant for the Responder's rejection threshold, where economics education does not seem to have a causal effect on our regression analysis. We believe that this outcome could potentially be caused by our small sample size in the social-preferences module, a limitation of our design. Alternatively, it also could be caused by the differ-

[21] We also restricted students to be in the common support of the rank-order list, limiting the treatment group to students who had at least one non-economics/business major in their preferences. Consequently, in Table B.8, the number of observations is smaller than 963.

[22] We do not plot a RD-type figure for the reciprocity specification because the parameter of interest concerns an interaction term (econ * $M'$).

ences between the two exercises: While the second exercise imposes fewer restrictions on functional forms, it does not include the controls on student-major preferences. Figure 2.7 presents our results on probabilistic beliefs, and visually the findings are in line with our regression estimates.

## 2.6 Conclusion

Identifying the effect of majors on decision-making is intricate, in part due to students' initial preferences and self-selection in their application. This paper takes advantage of a natural experiment in China to analyze the causal impact of an economics/business major on decision-making. We analyze data by matching survey results for college students near the economics-admission cutoffs to examine the effect an economics education has on peoples' decision-making along several dimensions.

The main results of our paper address three aspects. First, students who receive an economics education are more likely to behave as risk-neutral agents in small-stake choices. On the other hand, our finding that economics and non-economics students are equally sensitive to a loss frame in the Willingness to Pay (WTP) game suggests that education as a debiasing scheme is not guaranteed—even extensive and long-term training may not change some fundamental heuristics.

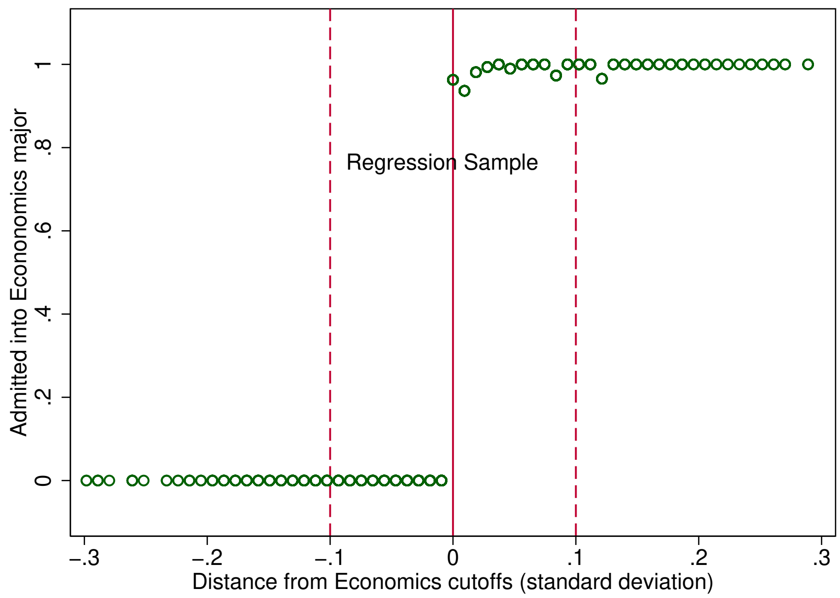
Second, an education in economics/business majors seems to shift individuals' first-order beliefs about others' social preferences (altruism, norms, etc.) more than it shifts students' own social preferences. Individuals who receive an economics education may believe that other people are homo-economicus. This finding may have significant implications for how people who receive an economics education might interact with others.

Third, courses in statistics, which are required in economics/business curricula, successfully endow people with correct probabilistic beliefs.

Taken together, our results show a mixed picture of exposure to economics education: On the one hand, these curricula improve students' decision-making qualities without altering their own social preferences; on the other hand, these curricula lead to substantially biased beliefs about other peoples' social preferences, which could impact social interaction in real life.

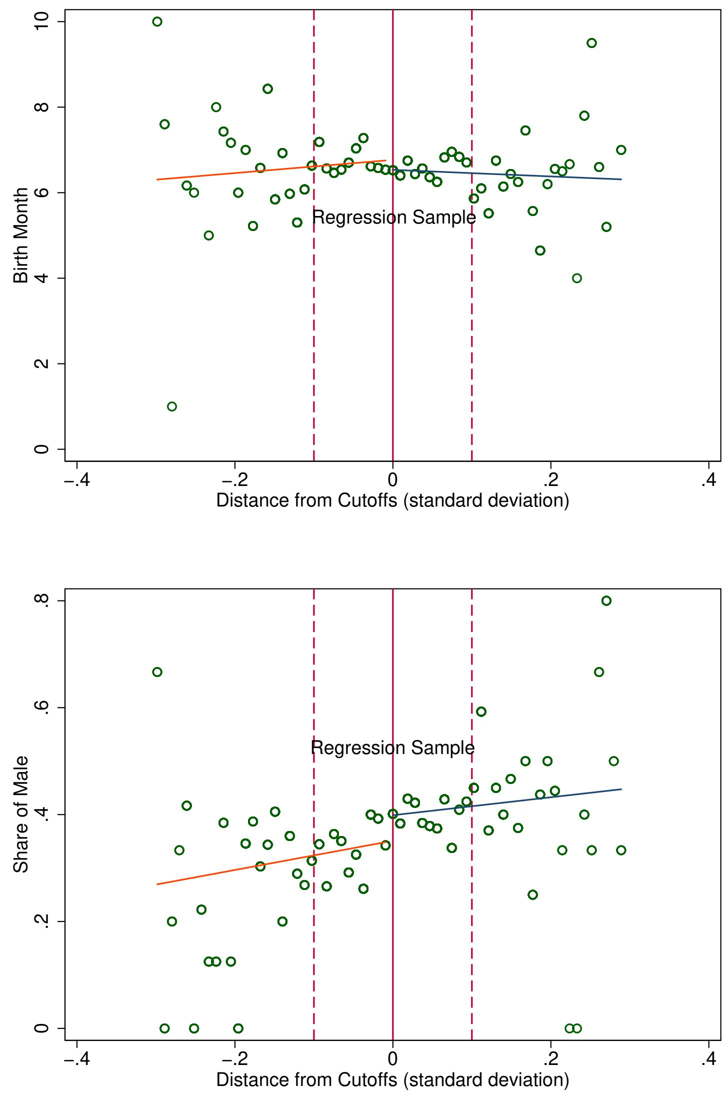
While we have found that the most plausible explanation for our results hinges on the effects of an economics education on students' decision-making, there are three limitations that impede us from drawing stronger conclusions. First, despite the systematic changes in students' survey responses, it remains relatively unclear how much such gaps could affect real-life actions. As we discuss in Section 2.4, students may treat the surveys as exam questions, and economics students might be more capable or exert more effort when responding to these economics-related questions. Second, despite the suggestive findings that within the sample, non-causal specification may overestimate the learning effect in several cases, strictly speaking, we cannot compare the magnitude of this learning effect to the full extent of sorting because most invited survey participants were those whose most preferred choices were economics/business majors. Third, the effects we discovered are limited to one particular university, and the estimates in social-preferences module are not very precise due to the limits of our design. We hope that future research can shed more light on the external validity of our findings and the effects of such preference gaps on real-life decision-making.

Figure 2.1: Probability of being Admitted to Economics and Distance to Cutoff



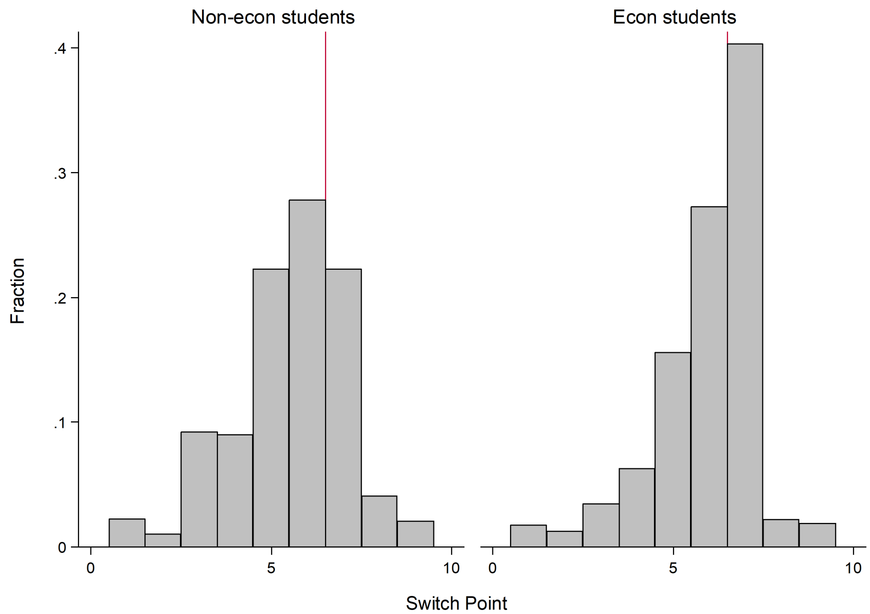

Note: This figure depicts the probability of being admitted to an economics/business major against the distance to the cutoff score using the administrative data of every student. The horizontal axis indicates the distance between an exam score and the corresponding threshold of an economics-admission line. The vertical axis denotes the probability of being admitted to an economics/business major.

Figure 2.2: Birth Month/ Gender Distribution and Distance to Cutoff Using the Administrative Data



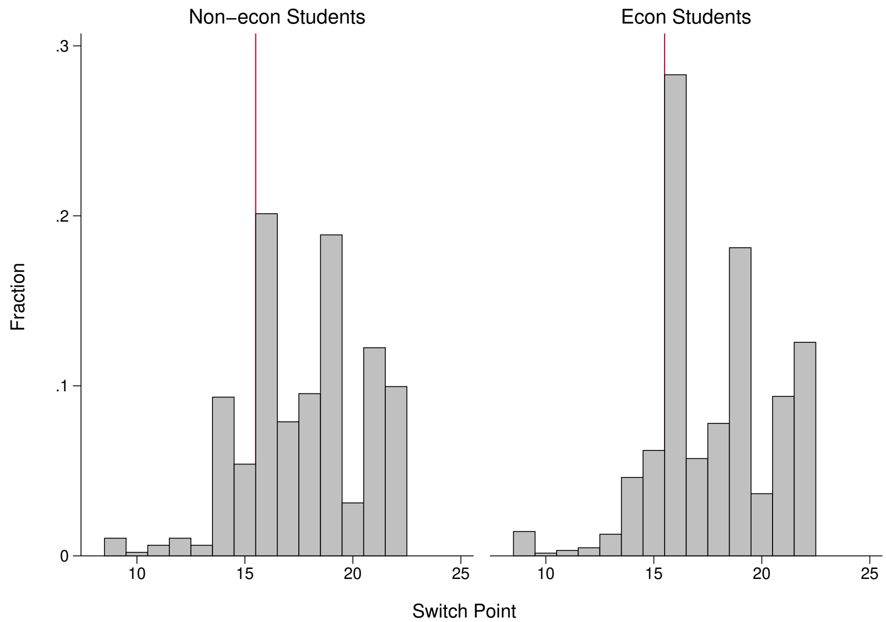

Note: This figure shows the graph of the distribution of pre-determined variables against the distance to the economics/business cutoff score. The vertical axis in the top panel shows the average birth month, and the vertical axis in the bottom panel denotes the share of males conditional on the distance to a cutoff score. The relationship between birth month/ gender distribution and the distance to an economics admission line shows that there is no systematic difference for people near the cutoff scores. The balance test results for the other control variables are shown in Table 2.2

Figure 2.3: Distribution of Switching Points for Risk: MPL 1



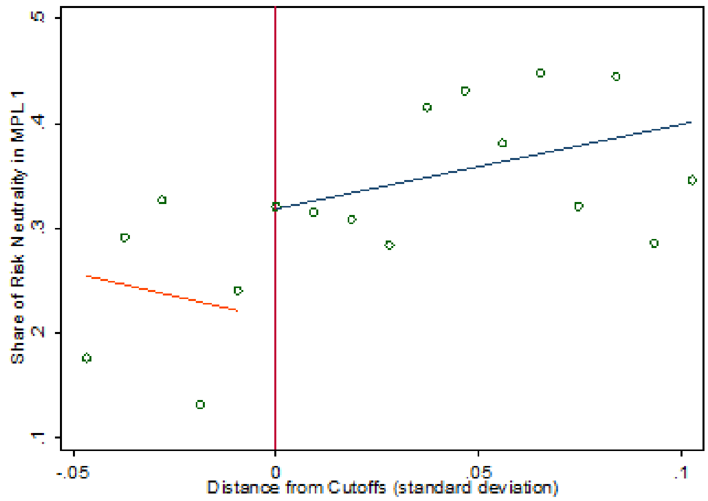

Note: This figure presents the distribution of students' choices in MPL 1, where the red line in each histogram indicates the switching point that indicates risk neutrality. Please see Table 2.3 for the coding of the switching point.

Figure 2.4: Distribution of Switching Points for Risk: MPL 2

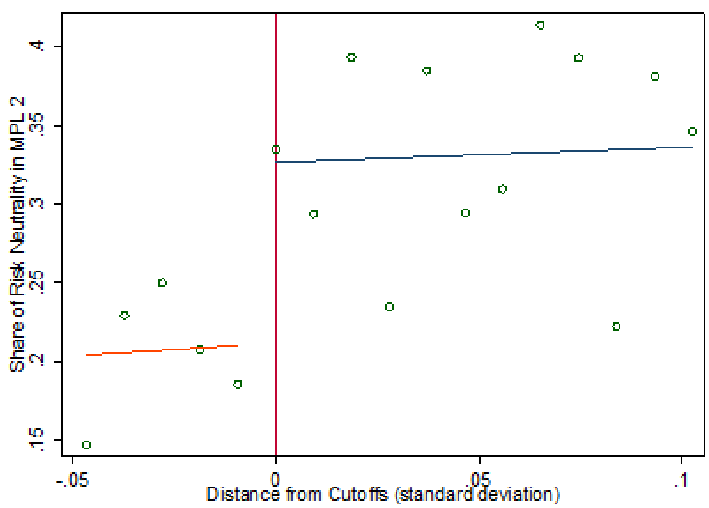



Note: This figure presents the distribution of students' choices in MPL 2. Please see Table 2.3 for the coding of the switching point.

Figure 2.5: Share of Risk Neutrality in MPL 1 and MPL 2 against Distance-to-cutoff



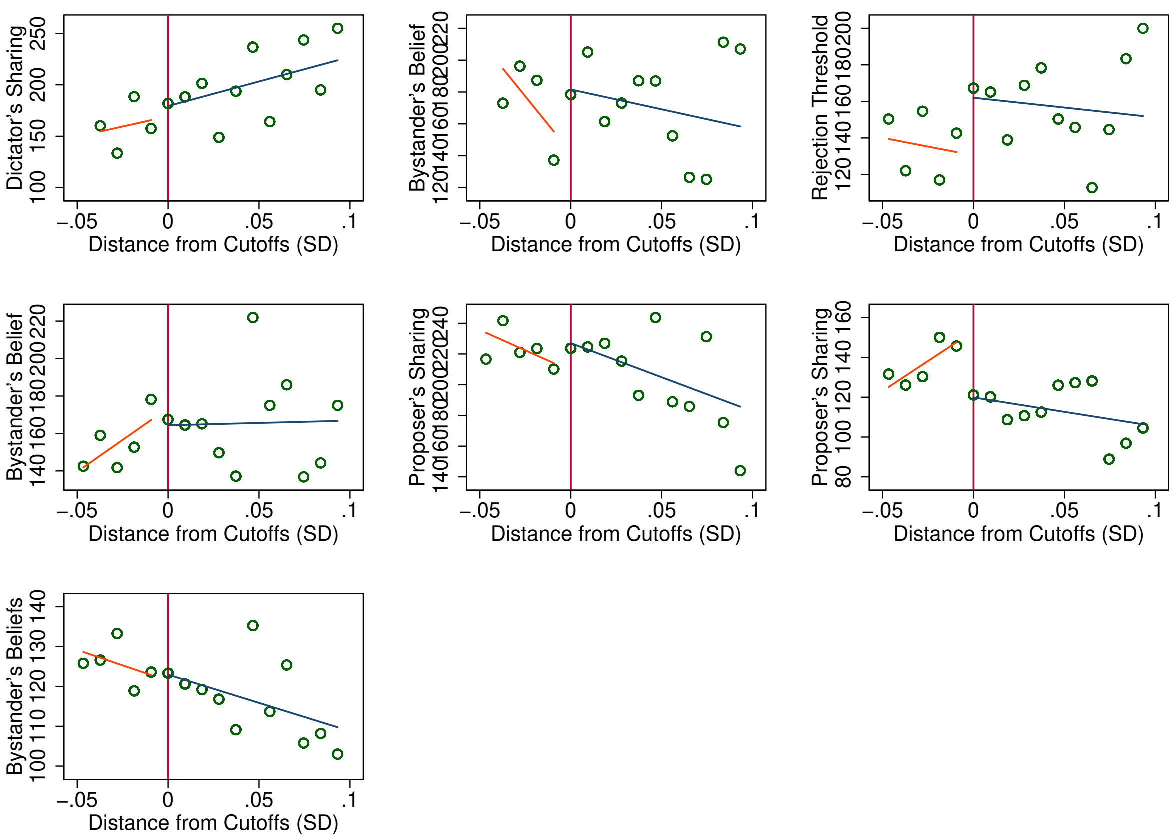

(a) Share of Risk Neutrality in MPL 1




(b) Share of Risk Neutrality in MPL 2

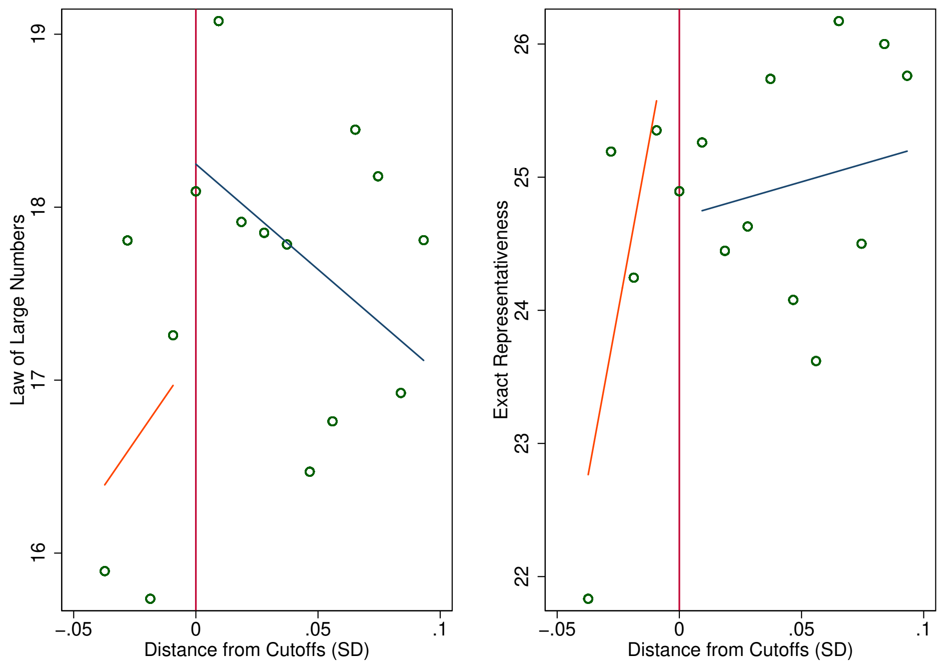
Note: This RD-type figure presents the share of risk neutrality in MPL 1 and MPL 2 against the distance-to-cutoff score of an economics/business major. Consistent with Table 2.4, the shares of risk neutrality are discontinuous around the cutoff.

Figure 2.6: Social Preferences in Three Games against Distance-to-cutoff




Note: This figure shows the outcome variables of the Dictator's Game, Ultimatum Bargaining Game, and Trust Game against the distance to an economics/business cutoff score. The first two plots of the top panel present the Dictator's sharing and Bystander's belief in the Dictator Game (Table 2.6). The third plot in the top panel and the first two plots in middle panel show the patterns of the Rejection Threshold, Bystander's belief and Proposer's sharing in the Ultimatum Bargaining Game (Table 2.7). And the third plot in the middle panel and the plot in the bottom panel display the patterns of the Proposer's sharing and Bystander's belief in the Trust games (Table 2.8).

Figure 2.7: Probabilistic Beliefs against Cutoff Scores




.

Note: This figure demonstrates the patterns of the law of large number and exact representativeness against the distance to an economics/business cutoff score (Table 2.9).

Table 2.1: Cutoff Construction and Identification Strategy: Four Examples

| Rank Order | Ranked Major | Field | Admission Cutoff |
|---|---|---|---|
| ***Panel A: Qian and Wu's ROL*** | | | |
| 1st best | Finance | Economics & Business | 635 |
| 2nd best | Accounting | Economics & Business | 631 |
| 3rd best | Industry Management | Economics & Business | 632 |
| 4th best | Chinese language & Literature | Humanity & Language | 629 |
| 5th best | Law | Law & Sociology | 628 |
| 6th best | Sociology | Law & Sociology | 616 |
| Qian's CEE Score = 631 | | Admitted Field: Economics & Business | |
| Wu's CEE Score = 629 | | Admitted Field: Humanity & Language | |
| ***Panel B: Lin and Wang's ROL*** | | | |
| 1st best | Finance | Economics & Business | 635 |
| 2nd best | Law | Sociology | 631 |
| 3rd best | Economics | Economics & Business | 629 |
| 4th best | Marketing | Economics & Business | 626 |
| 5th best | Journalism | Humanity & Language | 622 |
| 6th best | Sociology | Law & Sociology | 616 |
| Lin's CEE Score = 632 | | Admitted Field: Sociology | |
| Wang's CEE Score = 629 | | Admitted Field: Economics & Business | |

Table 2.2: Summary Statistics

| Variables | Non economics | Economics | Difference |
|---|---|---|---|
| Gender (Female=0, Male=1) | 0.36 | 0.38 | -0.02 |
| | | | (0.02) |
| Father's education | 13.63 | 13.79 | -0.16 |
| | | | (0.19) |
| Mother's education | 13.15 | 13.40 | -0.24 |
| | | | (0.21) |
| Pre-college ranking | 13.26 | 11.86 | 1.40* |
| | | | (0.83) |
| 6-Month consumption | 13304.60 | 13068.56 | 236.03 |
| | | | (709.81) |
| 6-Month allowance | 13139.96 | 12321.44 | 818.52 |
| | | | (786.36) |
| Years of boarding before college | 2.05 | 2.02 | 0.03 |
| | 496 | 493 | (0.13) |

This table presents the summary statistics of characteristics between economics/business and non-economics/business students. The first column shows the name of the variables for which we conduct the balance test.

Table 2.3: Interpretation of Choice for MPL 1 and MPL 2 under CRRA

| Choice | Interpretation | Switching point (question #) |
|---|---|---|
| Panel A: Multiple Price List 1 | | |
| Always choose certain payment | Dominated choice | 1 |
| Choose the fixed lottery iff $p = 25$ | Dominated choice | 2 |
| Choose the fixed lottery iff $p \leq 30$ | Risk averse | 3 |
| Choose the fixed lottery iff $p \leq 35$ | Risk averse | 4 |
| Choose the fixed lottery iff $p \leq 40$ | Risk averse | 5 |
| Choose the fixed lottery iff $p \leq 45$ | Risk averse | 6 |
| Choose the fixed lottery iff $p \leq 50$ | Risk neutral | 7 |
| Choose lottery iff $p \leq 55$ | Risk loving | 8 |
| Always choose the fixed lottery | dominated choice | 9 |
| Panel B: Multiple Price List 2 | | |
| Never choose B | Risk averse | 9 |
| Choose B iff $p = 0.25$ | Risk averse | 10 |
| Choose B iff $p \geq 0.23$ | Risk averse | 11 |
| Choose B iff $p \geq 0.21$ | Risk averse | 12 |
| Choose B iff $p \geq 0.19$ | Risk averse | 13 |
| Choose B iff $p \geq 0.17$ | Risk averse | 14 |
| Choose B iff $p \geq 0.15$ | Risk neutral | 15 |
| Choose B iff $p \geq 0.13$ | Risk neutral | 16 |
| Choose B iff $p \geq 0.11$ | Risk loving | 17 |
| Choose B iff $p \geq 0.09$ | Risk loving | 18 |
| Choose B iff $p \geq 0.07$ | Risk loving | 19 |
| Choose B iff $p \geq 0.05$ | Risk loving | 20 |
| Choose B iff $p \geq 0.03$ | Risk loving | 21 |
| Always choose B | Risk loving | 22 |

This table presents the interpretation of choices for MPL 1 and MPL 2 under the assumption that students have CRRA preferences. The second column shows the interpretation (risk averse/neutral/loving) corresponding to each potential choice in the first column. The third column presents the encoding of the responses in the first column.

Table 2.4: Risk Preference and Distribution in MPL 1 and MPL 2

| | MPL 1 (1) - (3) | | | MPL 2 (4) - (6) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dep. var. | Risk Neutral | Risk Loving | Risk Averse | Risk Neutral | Risk Loving | Risk Averse |
| Econ=1 | 0.118*** | -0.019 | -0.099*** | 0.060** | -0.051* | 0.003 |
| | (0.035) | (0.015) | (0.036) | (0.031) | (0.027) | (0.036) |
| Constant | 0.032** | 0.034** | 0.800*** | 0.150*** | 0.035* | 0.816*** |
| | (0.039) | (0.017) | (0.041) | (0.044) | (0.019) | (0.046) |
| Major-Preference FX | X | X | X | X | X | X |
| Inconsistent Choice Excluded | X | X | X | X | X | X |
| Common Support of Major Preference | X | X | X | X | X | X |
| Observations | 765 | 765 | 765 | 813 | 813 | 813 |
| R-squared | 0.041 | 0.013 | 0.031 | 0.039 | 0.010 | 0.030 |
| Mean outcome of non-econ | 0.23 | 0.04 | 0.72 | 0.20 | 0.18 | 0.57 |
| | (0.42) | (0.21) | (0.45) | (0.40) | (0.39) | (0.50) |

This table reports the regression results using equation (1) for risk preferences in MPL 1 (column (1)-(3)) and MPL 2 (column (4)-(6)). The leftmost column is the name of our key variables added to the regression. *Econ* is taking the value of one if a students was in an economics/business related major. Columns (1) and (4) estimate the impact of an economics/business major (Econ=1) on the share of students who appear to be risk neutral. Columns (2) and (5) estimate the impact of an economics/business major on the share of risk-loving students. Columns (3) and (6) estimate the impact of an economics/business major on the proportion of risk-averse students.
All Columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics/business major category (*Major-Preference FX*). In these columns, we also exclude individuals with inconsistent choices, and limit to students who put both economics and non-economics/business majors in the rank-order list (*Common Support of Major Preference*). Columns (1)-(3) additionally exclude people who make dominated choices.
Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 2.5: Risk Preference and Distribution in MPL 1

| Dep. var. | (1) Risk Neutral WTA | (2) Risk Neutral WTP | (3) Risk Loving WTA | (4) Risk Loving WTP | (5) Risk Averse WTA | (6) Risk Averse WTP |
|---|---|---|---|---|---|---|
| Econ=1 | 0.184*** | 0.027 | -0.034 | -0.014 | -0.150*** | -0.013 |
| | (0.051) | (0.047) | (0.027) | (0.012) | (0.054) | (0.048) |
| Constant | 0.199*** | 0.131** | 0.069** | 0.003 | 0.732*** | 0.865*** |
| | (0.059) | (0.052) | (0.031) | (0.014) | (0.062) | (0.053) |
| Major-Preference FX | X | X | X | X | X | X |
| Inconsistent Choice Excluded | X | X | X | X | X | X |
| Common Support of Major Preference | X | X | X | X | X | X |
| Observations | 374 | 391 | 374 | 391 | 374 | 391 |
| R-squared | 0.079 | 0.032 | 0.020 | 0.014 | 0.048 | 0.037 |

This table reports the regression results for risk preferences using Willingness to Pay/Willingness to Accept (WTP/WTA) subsamples. Columns (1), (3), and (5) describe the treatment effects of receiving economics/business education on subjects who make choices under WTA framing. Columns (2), (4), and (6) detail the treatment effects on subjects who make choices under WTP framing. Columns (1) and (2) estimate the impact of an economics/business education ($Econ = 1$) on the share of students who appear to be risk neutral. Columns (3) and (4) estimate the impact of an economics/business education on the share of risk-loving students. Columns (5) and (6) estimate the impact of an economics/business education on the share of risk-averse students.
All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics/business major category (*Major-Preference FX*), and exclude people who make dominated or inconsistent choices. We additionally limit the regression sample to students who put both economics and non-economics/business majors in their rank-order list (*Common Support of Major Preference*).
Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 2.6: Social Preferences in Dictator Game

| Dep. Var | (1)<br>Dictator's Sharing | (2)<br>Dictator's Sharing | (3)<br>Bystander's Belief | (4)<br>Bystander's Belief |
|---|---|---|---|---|
| Econ=1 | 2.591 | 0.239 | -26.550* | -22.990* |
| | (14.181) | (13.885) | (14.080) | (13.861) |
| Common Support of Major Preference | | X | | X |
| Observations | 335 | 274 | 344 | 275 |
| R-squared | 0.036 | 0.050 | 0.018 | 0.018 |
| Mean of Non-Econ | 186.42 | 187.87 | 174.31 | 179.19 |
| | (104.89) | (101.95) | (112.54) | (106.11) |

This table presents the regression results using equation (1) for social preferences in the Dictator Game. The dependent variable is the Dictator's sharing in columns (1) and (2), and the Bystander's belief regarding the mean of Dictator sharing in columns (3) and (4).
All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics/business major category (*Major-Preference FX*). In columns (2) and (4), we additionally limit the regression sample to students who put both economics and non-economics/business majors in the rank-order list (*Common Support of Major Preference*).
Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 2.7: Social Preferences in the Ultimatum Game

| Dep. Var. | (1)<br>Rejection Threshold | (2)<br>Rejection Threshold | (3)<br>Bystander's Belief | (4)<br>Bystander's Belief | (5)<br>Proposer's Sharing | (6)<br>Proposer's Sharing |
|---|---|---|---|---|---|---|
| Econ=1 | 9.694 | 13.851 | 0.710 | -4.101 | -11.512 | -10.370 |
| | (12.303) | (12.574) | (12.546) | (12.820) | (10.269) | (9.587) |
| Mean of | 147.18 | 147.18 | 160.96 | 147.18 | 228.12 | 228.12 |
| Non-econ | ( 87.41) | (87.40) | (87.66) | ( 87.41) | (62.51) | (62.51) |
| Common Support of Major Preference | | X | | X | | X |
| Observations | 336 | 274 | 316 | 262 | 337 | 266 |
| R-squared | 0.018 | 0.027 | 0.012 | 0.014 | 0.041 | 0.039 |

This table presents the regression results using equation (1) for social preferences in the Ultimatum Game. The dependent variable is the Rejection Threshold of Player B in columns (1) and (2), the Bystanders' belief regarding the mean amount of Player A's sharing in columns (3) and (4), the actual mean of Player A's sharing in columns (5) and (6).
All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics/business major category (*Major-Preference FX*). In columns (2), (4), and (6), we additionally limit the regression sample to students who put both an economics and non-economics/business major in their rank-order list (*Common Support of Major Preference*).
Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 2.8: Social Preferences in the Trust Game

| VARIABLES | (1) Proposer's Sharing | (2) Bystander's Belief | (3) Reciprocity of Player 2 | (4) Bystander's Belief about Reciprocity |
|---|---|---|---|---|
| Econ* M' (-50, 0, 50) | | | -0.031 | -0.383** |
| | | | (0.090) | (0.154) |
| Econ=1 | -13.575** | 0.122 | -0.992 | -1.114 |
| | (6.541) | (6.428) | (3.905) | (6.794) |
| M'=(-50, 0, 50) | | | 1.201*** | 1.151*** |
| | | | (0.056) | (0.098) |
| Common Support of Major Preference | X | X | X | X |
| Mean of Dependent Variable | 122.13 | 121.92 | 97.78 | 97.62 |
| Observations | 409 | 393 | 801 | 801 |
| R-squared | 0.020 | 0.007 | 0.486 | 0.190 |

This table presents the regression results using equation (2) for social preferences in the Trust Game. Column (1) analyzes how an economics education affects students' sharing behavior as a Proposer in the Trust Game, which could be interpreted as students' beliefs regarding the amount that the other players would like to reciprocate. The dependent variable is Bystanders' belief regarding the mean amount of Player A's sharing in the Trust Game in column (2). Columns (3) and (4) ask Player B the amount they would like to give back if Player A gives her 50, 100, 150 RMB and the Bystander's belief regarding the mean amount of Player B's giving back, respectively.
All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics/business major category (*Major-Preference FX*), and limit the regression sample to students who put both economics and non-economics/business major in their rank-order list (*Common Support of Major Preference*).
Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 2.9: Probabilistic Beliefs

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. Var.: Coins on asset B | Law of Large Numbers | | Indifferent Choices | | Exact Representativeness | |
| Econ=1 | 0.995** | 1.219*** | -0.251 | -0.331 | 0.952** | 1.134** |
| | (0.457) | (0.462) | (0.362) | (0.379) | (0.481) | (0.474) |
| Major-Preference FX | X | X | X | X | X | X |
| Common Support of Major Preference | | X | | X | | X |
| Constant | 17.102*** | 17.092*** | 15.409*** | 15.413*** | 24.299*** | 24.305*** |
| | (0.297) | (0.273) | (0.235) | (0.224) | (0.313) | (0.280) |
| Observations | 989 | 802 | 989 | 802 | 989 | 802 |
| R-squared | 0.013 | 0.022 | 0.006 | 0.008 | 0.015 | 0.023 |

This table presents the results using equation (1) for probabilistic belief outcomes. Column (1) reports the treatment effect of an economics education on question 1 for the probabilistic beliefs (testing their knowledge of the law of large numbers). Column (2) reports the treatment effects on question 2 for the probabilistic beliefs where no psychological heuristics are linked to this question and any allocation is optimal. Column (3) reports the treatment effects on question 3 for the probabilistic beliefs (testing knowledge of exact representativeness).
All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics/business major category (*Major-Preference FX*). In columns (2), (4) and (6), we additionally limit the regression sample to students who put both economics/business and non-economics/business majors in their rank-order list (*Common Support of Major Preference*).
Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 2.10: Exposure and Learning Effects

| Dep. Var. | (1) MPL1 Risk Neutral | (2) MPL2 Risk Neutral | (3) Law of Large Numbers | (4) Identical Choices | (5) Exact Representiveness | (6) Dictator's Sharing | (7) Bystander's Belief |
|---|---|---|---|---|---|---|---|
| Econ*Freshman | 0.095* | 0.051 | 0.131 | -0.529 | 0.667 | 17.980 | 8.269 |
| | (0.051) | (0.052) | (0.679) | (0.558) | (0.699) | (19.408) | (20.388) |
| Econ*Post Freshman | 0.177*** | 0.084** | 1.991*** | -0.512 | 1.680*** | -19.103 | -27.888* |
| | (0.036) | (0.037) | (0.480) | (0.395) | (0.494) | (13.998) | (14.619) |
| Common Support of Major Preference | X | X | X | X | X | X | X |
| Constant | 0.220*** | 0.252*** | 17.004*** | 15.484*** | 24.210*** | 191.141*** | 185.934*** |
| | (0.020) | (0.020) | (0.265) | (0.218) | (0.273) | (7.963) | (8.175) |
| Observations | 802 | 802 | 802 | 802 | 802 | 274 | 275 |
| R-squared | 0.030 | 0.007 | 0.022 | 0.003 | 0.014 | 0.013 | 0.016 |

This table examines the heterogeneity of treatment effects with regard to exposure to an economics education. It presents the average differences between economics students (in a particular year) and non-economics students in the responses on risk preferences, probabilistic beliefs and social preferences, introducing two interaction terms of $Econ \times Freshman$ and $Econ \times Post\_Freshman$ using equation (3). The outcome variable in Columns (1) and (2) is the share of risk-neutral students in MPL 1 and MPL 2. Columns (3), (4), and (5) pertain to the probabilistic belief questions on the law of large numbers (LLN), two identical choices, and the probabilistic belief questions on Exact Representativeness (ER), respectively. The outcome variables in columns (6) and (7) are the Dictator's actual sharing and Bystander's beliefs regarding Dictator's sharing in the Dictator Game.

All columns limit the regression sample to students who put both economics and non-economics/business major in their rank-order list (*Common Support of Major Preference*).

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

## Table 2.11: Non-Causal Effects v.s. Causal Effects of Economics Education

| | Risk Preferences (1)-(2) | | Probabilistic Beliefs (3)-(5) | | | Dictator Game (6)-(7) | | Trust Game (8)-(11) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| | MPL1 | MPL2 | Law of Large | Two Indifferent | Exact | Dictator's | Bystander's | Proposer's | Bystander's | Reciprocity of | Bystander's Belief |
| Dep. Var. | Risk Neutral | Risk Neutral | Numbers | Choices | Representiveness | Sharing | Belief | Sharing | Belief | Player 2 | about Reciprocity |
| **Panel A: Non-causal Regression: Non-Gaokao Sample Included** | | | | | | | | | | | |
| Econ=1 | 0.153*** | 0.053** | 1.324*** | -0.104 | 1.729*** | -16.217* | -13.950 | -7.686* | 5.547 | 7.052 | -2.636 |
| | (0.023) | (0.023) | (0.301) | (0.249) | (0.323) | (9.120) | (9.850) | (4.273) | (4.980) | (12.695) | (3.242) |
| M′=(-50, 0, 50) | | | | | | | | | | 1.199*** | |
| | | | | | | | | | | (0.246) | |
| Econ*M′(-50, 0, 50) | | | | | | | | | | 0.258 | |
| | | | | | | | | | | (0.315) | |
| M′=(-50, 0, 50) | | | | | | | | | | | 1.054*** |
| | | | | | | | | | | | (0.063) |
| Econ*M′(-50, 0, 50) | | | | | | | | | | | -0.018 |
| | | | | | | | | | | | (0.079) |
| Observations | 1,634 | 1,634 | 1,634 | 1,634 | 1,634 | 551 | 528 | 825 | 806 | 1,631 | 1,630 |
| R-squared | 0.026 | 0.003 | 0.012 | 0.000 | 0.017 | 0.006 | 0.004 | 0.004 | 0.002 | 0.046 | 0.313 |
| **Panel B: Non-causal Regression: No. of Stated Preferences Using Non-causal Sample** | | | | | | | | | | | |
| # Preferred Econ Majors | 0.031*** | 0.017** | 0.247*** | -0.015 | 0.437*** | -5.698** | -2.171 | -2.576** | -2.507* | -5.946 | -0.510 |
| among Six Preferences | (0.007) | (0.006) | (0.085) | (0.070) | (0.091) | (2.557) | (2.842) | (1.212) | (1.393) | (3.662) | (1.102) |
| Observations | 1,634 | 1,634 | 1,634 | 1,634 | 1,634 | 551 | 528 | 825 | 806 | 1,631 | 1,630 |
| R-squared | 0.013 | 0.004 | 0.005 | 0.000 | 0.014 | 0.009 | 0.001 | 0.005 | 0.004 | 0.002 | 0.000 |
| **Panel C: Causal Regression: Table 2.4,2.6,2.8, 2.9** | | | | | | | | | | | |
| Econ=1 | 0.118*** | 0.060** | 1.219*** | -0.331 | 1.134** | 0.239 | -26.550* | -13.575** | 0.122 | -0.992 | -1.114 |
| | (0.035) | (0.031) | (0.462) | (0.379) | (0.474) | (13.885) | (14.080) | (6.541) | (6.428) | (3.905) | (6.794) |
| M′=(-50, 0, 50) | | | | | | | | | | 1.201*** | |
| | | | | | | | | | | (0.056) | |
| Econ*M′(-50, 0, 50) | | | | | | | | | | -0.031 | |
| | | | | | | | | | | (0.090) | |
| M′=(-50, 0, 50) | | | | | | | | | | | 1.151*** |
| | | | | | | | | | | | (0.098) |
| Econ*M′(-50, 0, 50) | | | | | | | | | | | -0.383** |
| | | | | | | | | | | | (0.154) |

This table reports the regression results using a non-causal sample (panel A and B) and the causal regression sample used by this paper (panel C). The dependent variables in columns (1) and (2) are the share of risk neutral students in MPL 1 and MPL 2, in which we pool sample from both Willingness to Accept (WTA) and Willingness to Pay (WTP) mode. Columns (3), (4), and (5) report results on the law of large numbers (LLN), two indifferent choice question, and the probabilistic belief question on Exact Representativeness (ER), respectively. The outcome variables in columns (6) and (7) are the Dictator's actual sharing and Bystander's beliefs regarding the Dictator's sharing in the Dictator Game. Column (8) analyzes how an economics education affects students' sharing behavior as a Proposer in the Trust Game, which could be interpreted as students' beliefs regarding the amount that the other players would like to reciprocate. The dependent variable is Bystanders' beliefs regarding the mean amount of Player A's sharing in the Trust Game in column (9). Columns (10) and (11) ask Player B the amount they would like to give back if Player A gives 50, 100, 150 Yuan and Bystander's beliefs regarding the mean amount of Player B's giving back.

Panel B reports the results using regression specification (4), # *Preferred Econ Majors* represents how many economics/business majors a student $i$ put in her ROL.

Panel C copies the previous coefficients in Section 2.4 for comparison. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

# Chapter 3

# Imperfect Land Market, Migration Cost, and Resources Reallocation: Evidence from China

## 3.1 Introduction

Under a complete land and labor market, the household demand for farm labor and individual occupational choice are perfectly separable from the initial land endowment (Benjamin, 1992; LaFave and Thomas, 2016). However, developing countries are often characterized by barriers to productivity growth and structural transformation that act at the market level, such as financial access or labor mobility constraints. In this setting, a central research question is how are factors, like the initial endowments, allocated? Economics theories state that the optimal allocation requires well-functioning markets, contract mechanisms, and an environment to enforce contracts. A canonical example is the land market in developing countries. However, such factor allocation problem is often ignored in the literature, e.g., how to price land and labor when no market exists? Most literature argues that land fragmentation increases with population growth/ each generation. However, little has been done to demonstrate how land is divided or allocated when land transactions are extremely uncommon. Are the effects of the initial allocation of land on households' efficiency with or without a market heterogeneous across different land sizes?

We investigate these issues and provide novel evidence on how heterogeneity in initial land endowments caused by family-size-based land reallocation affects household and individual short- and long-run efficiency without markets. We also study how the induced distortion in factor inputs evolves with the formation of the labor market and land contract system. Under a communal land system, which occasionally reallocated land at the village level based on family size (Zhao, 2020), we construct variations in household initial land size through tracking household size change before the last redistribution prior to 2003. The Rural Land Contract Law issued by the central government in 2003 prohibited any land redistribution and created an unexpected "last reallocation" for each village, which was barely predicted by each household within the village. If households had a family size change just before the "last reallocation", their land endowments would be totally different from households with a family size change just after the "last reallocation". We additionally introduce prefecture-level variations in rural labor market reform and provincial-level variations in the enforcement of land contract laws[1] to learn about the mechanisms and consequences of the market formation. Our main findings are that households oversupplied labor in agriculture when markets were missing. These distortions and misallocations in labor and land were alleviated once rural-urban migration restrictions were relaxed and contracting laws in land renting activities were established.

We start by describing several stylized facts about China's land and rural labor market, motivating our inquiry into this study. 1) During 1982 and 2003, village land rights were occasionally redistributed between households based on family size, while no land contract existed, and market transactions among individual households were not allowed. Consequently, households had incentives to oversupply labor on the farm to reserve and expect more land for the subsequent reallocation. In the meantime, the rural labor market was incomplete due to the household registration system (Hukou system), which prohibited/restricted rural labor from migrating to and working in the urban area. As a result, the individual occupational choice was limited to either farm labor or non-agricultural job within the village. 2) These constraints were removed between 2003 and 2019 with the central government's support. The implementation of the land market and contracts was initiated in 2003, the Rural Land Contracting Law (RLCL), which gave farmers legal rights to lease their land while re-iterating existing protections for the security of land rights. 3) During this period, the rural-urban migration channel was also staggeringly relaxed due to the Hukou reform, which granted work permits in the urban area for rural households. The land reform has spurred a 10% increase in land renting activities,[2] while the reduction in migration costs driven by Hukou reform raised the intra- and inter-provincial migrant population by 15 and 81% respectively Tombe and Zhu (2019).

To evaluate the heterogeneous treatment effects of initial land endowments in response to RLCL and Hukou reform, we collected novel data on the prefecture-level timing of the first wave of the Hukou reform, which encouraged an integrated labor market between rural and the urban area. Building on work by Chari et al. (2021), we also exploit the implementation of the central land contract law, announced in 2003. We combine the prefecture-level Hukou reform and province-level implementation data with a nationally representative panel dataset of 20,000 households in 360 villages between 1986 and 2015 (the National Fixed Point Survey or NFP). It includes 1) a panel of village surveys that provides information on agricultural output, land lease and use, income, and expense; 2) a panel of household surveys that provides information on landholding, input, output

[1] See the same variations in land reform in Chari et al. (2021).

[2] Chari et al. (2021) find that a 1.4 to 1.5 percentage point increase in the probability of engaging in land rentals in villages after the implementation of the reform, corresponding to a 10% increase.

consumption; 3) a panel of individual survey provides information on born year, gender, educational attainment, disability status. To find the village-level timing of the last land reallocation before 2003, we use supplementary data from the Village Democracy Survey (VDS) built by Martinez-Bravo et al. (2017), which records village administrative data and includes information about the timing of each round of reallocation for 183 villages.

We begin our novel analysis with RD-type sample construction and define two treatment and two control groups, respectively: 1) $T_1$: households that experienced an increase in family size before the last reallocation; (2) $C_1$: households that experienced an increase in family size right after the last reallocation; (3) $T_2$: households that experienced a decrease in family size before the last reallocation; (4) $C_2$: households that experienced a size reduction right after the last reallocation. Comparing the differentiated evolution path between $T_1$ versus $C_1$ and $T_2$ versus $C_2$ sheds light on the productivity-land-match efficiency under a system with missing markets. We then employ difference-in-difference estimates of the impact of the Hukou reform. The estimates exploit the staggered timing of implementation across prefectures and the effects of RLCL reform that gave farmers legal rights to rent the land. We are particularly interested in understanding whether the formation of the rental and labor market increases or decreases the productivity-land-match efficiency compared to the case without markets. One possibility, in theory, is that the markets might mitigate the misallocation since households with high productivity but low endowments can now rent more lands. However, the theory also predicts that high-productivity households with high endowments might rent out their lands and work in non-agricultural sectors as migrants once they have access to the rental and labor market. Thus, it remains puzzling whether the transition from a missing market to a complete market exacerbates or relieves the distortions mentioned above.

We then develop a theoretical framework that captures important institutional features of the rural land market in China. Farmers have insecure tenure overland. They choose whether to cultivate on their farms, work in a non-agricultural job within the village, or work in the urban area as migrants. Occupational switching costs depend on the status of the land and labor market. Before the Hukou system's reform, the cost of rural-to-urban migration was sufficiently high, impeding farmers' migration. For land endowments, there are two scenarios in our model. The land reallocation scheme captures one scenario, and there is no formal land rental market. The other is a new regime with no land reallocation, and people are allowed to rent in/out their land. Before the land contract reform, current uncultivated land was subject to expropriation risk. After the reform, such probability drops drastically to a much lower level. As a result, before land reform, households oversupplied labor on the farm, causing low TFP and misallocation of labor. Land market formation and labor market reform will alleviate such distortions and reduce labor misallocation.

Consistent with model predictions, we find evidence that households with less land per capita rent more land prior to land and labor market reforms. Households experiencing family size increase before the last reallocation (group $T_1$) acquired 10% more land per capita, and 12% more per working labor, than households experiencing family size increase after the last reallocation (group $C_1$). Similarly, households that experienced family size reduction before the last reallocation (group $T_2$) received 10% less land per capita and 9.4% less land per working labor than households that experienced family size drop after the last reallocation (group $C_2$).

We then introduce the interaction term between RLCL reform and defined treatment groups, $T_1$ and $T_2$, respectively. By comparing the family size increase group ($T_1$ and $C_1$), we find that the land reform significantly granted more land to the $C_1$ group by 7% per working labor. At the same time, it did not influence households' labor allocation on the farm. Conversely, by comparing the family size decrease group ($T_2$ and $C_2$), we find that the land reform neither affected land per working labor nor changed households' labor allocation on the farm. Our interpretation behind such differentiated effects is that most households oversupplied labor in agriculture, which pulled down the marginal product of labor away from the optimum. In particular, the control $C_1$ groups experiencing a family size increase, but the increase is after the reallocation, now can rent in more land without adding labor inputs since the marginal product of labor was below the optimum. Households in the decrease group did not have incentives to rent out land or change labor input since the treatment groups $T_2$ and control $C_2$ groups that experienced a family size decrease had a higher marginal product of labor now with the current level of landholdings.

We also find a more salient effect of labor market reform on household labor allocation and individual occupational choice than the rural land contracting reform. After the local Hukou reform, households that experienced a family size increase after the reallocation, but were not allocated land accordingly ($C_1$), had a 73% lower total labor inputs than the previous period, corresponding to a 47% decrease in labor input per unit of land. We do not observe such an effect in the groups that increased family size before the reallocation ($T_1$). In contrast, both households that experienced a family size drop before and after the reallocation had a 47% - 51% reduction in labor input per unit of land ($T_2$ and $C_2$). We then pool Hukou reform and land reform together to study the simultaneous influence of the two market reforms. Surprisingly, unlike previous literature, which has focused on single land markets,[3] our empirical results show that labor market reform had a remarkable dominance over land reform in rural labor allocation and individual occupational choice. When adding Hukou reform before the land reform, the effect of land reform on labor allocation is almost neglectable.

Our work builds on and contributes to three main literature. Our research questions are mostly related to the work on agricultural productivity literature emphasizing misallocation and structure transformation, particularly that relating to aggregate impacts of market frictions with structural models (Adamopoulos et al., 2017; Chen, Restuccia, and Santaeulàlia-Llopis, 2021; Manysheva, 2021; Adamopoulos et al., 2022). A key challenge in this literature is a reliable calibration of the evolution of market frictions over time. Existing research mainly relies on cross-sectional moments using observational data, including information on labor supply, farm production, and incomes. Little has been done to exploit exogenous variations in the labor and land market and use these causal

[3]See Adamopoulos et al. (2017); Chari et al. (2021); Adamopoulos et al. (2022).

estimations to capture the market frictions. An exception is a reduced-form work from Chari et al. (2021), which use the staggered timing of land contracting laws to estimate the friction in the land market. We follow both strands of the literature and mitigate the divide between structural and reduced-form estimation. Our contribution to the literature is that we use multiple variations from the land market for identification rather than only using single variations from the land and labor market. To the best of our knowledge, this is the first paper to estimate the key parameters for frictions in the structural model using simultaneous and exogenous variations from multiple market failures, which haven't been explored.

Our findings also add to literature linking farm size more generally and productivity difference (Adamopoulos and Restuccia, 2014). Foster and Rosenzweig (2022) explains the U-shaped relationship between farm productivity and farm scale from two factors, including Transaction costs and economies of scale. Much emphasis has been put on the role of the insecure land tenure system and inefficient investment (Jacoby, Li, and Rozelle, 2002; Zhao, 2020), and land rights and factor reallocations (Banerjee, Gertler, and Ghatak, 2002; Field, 2007; Goldstein and Udry, 2008; De Janvry et al., 2015; Agyei-Holmes et al., 2020). These papers focus on estimating the farm size distribution and the magnitude of market frictions using reduced-form or structural approaches. Using Chinese data, We show that the extent to which market failures affect labor allocation and individual occupation choice is indeed heterogeneous and varies systematically with different initial land endowments. The initial land endowment plays an important role in shaping household short/long run production efficiency before and after the formation of multiple markets. In this strand of literature, measurement error also a concern (Bils, Klenow, and Ruane, 2021; Gollin and Udry, 2021; Aragón, Restuccia, and Rud, 2022). Our methodology can address such issues using a quasi-randomized land redistribution scheme combined with family size change, as well as two policy shocks.

This paper also relates to a large literature seeking to understand the barriers to reallocation of labor out of the agricultural sector in developing countries. Quantitative work in this area includes misallocation and corresponding productivity loss due to migration frictions (Bryan and Morten, 2019; Tombe and Zhu, 2019), rural-urban migrant workers, and firm productivity growth (Imbert et al., 2018), and rural-urban migration, sorting, and productivity gap (Hamory et al., 2021*b*; Gai et al., 2021). Several studies have documented that the Hukou system of local registration is particularly important in preventing worker sorting across space in China. Ngai, Pissarides, and Wang (2019), for instance, show that the Hukou leads to labor over-supply in agriculture and slows urbanization and industrialization. Our goal is to help policymakers decide which markets to prioritize.

The paper proceeds as follows. Section 3.2 provides detailed institutional background on land tenure laws and reform, as well as labor market reform in China. Section section 3.3 presents a simple theoretical model that provides intuition for the empirical analysis. Section 3.4 describes the data and our novel identification strategy. Section 3.5 and 3.6 present the empirical analysis, and Section 3.7 concludes.

## 3.2 Background

### Land Redistribution Scheme

Under the Household Responsibility System, instituted in the early 1980's, individual farming households were granted the right to use the land, carry out their production independently, and keep the residual income from farm activities. Households can decide what, how, and when to plant. However, the ownership of all rural land remains with rural collectives, which entitles village officials to periodically redistribute land among households within the same village for the egalitarian purpose.[4] Prior to 2003, pertaining to missing land markets, such a redistribution scheme could maintain the egalitarian distribution of land plots and eliminate the inefficiency caused by household-level demographic change, such as new household formation or exit of the existing households. In addition, village leaders could also use a land redistribution scheme as a "carrot and stick" to accomplish output quotas and collect taxes (Rozelle and Li, 1998).

Rural households need to hold several requirements to be eligible for the land redistribution: 1) they cannot convert agricultural land to other usages; 2) they cannot leave plots uncultivated for more than two years. Since each round of redistribution was based on previous family size, households had incentives to keep more members working on the farm to obtain more land, which caused great labor oversupply in agriculture without well-functioning labor and land market. Figure 3.1 presents the number of villages that experienced redistribution each year. We could observe that among 183 villages in our sample, eight villages had a redistribution in 1980, while this number was 52 in 1990 and 15 in 2002. After the implementation of the 2003 Rural Land Contract Law, abolishing the redistribution and forming the formal contracts and land rental markets, the frequency of redistribution dropped substantially to less than ten villages.[5] These frequent reallocations had slowly exacerbated the fragmentation of farmers' land plots,[6] which significantly discouraged agricultural investments.

### Labor Market Reform

During the period between 1982 and 2003, in conjunction with the land redistribution scheme, the labor market was also not well functioning, which restricted the mobility of migrant workers from rural to the urban area due to the Hukou system. Introduced in 1958, the Hukou is a system of local registration that severely restricts labor mobility in China, particularly rural farming labor mobility. Before 1978, workers were prohibited from working outside their region/registration category. Since then, the restrictions have been relaxed in incremental steps, with reforms being introduced gradually across

[4] See the official website from the Ministry of Agriculture and Rural Affairs of China for more details: http://www.moa.gov.cn/ztzl/gzypxjl/gzjl/201502/t20150204_4396141.htm

[5] Consistently, in data collected by Chari et al. (2021) in 2012, only 2.5% of households had experienced a government land reallocation in the past five years.

[6] Wu et al. (2018) shows that 98% of farming households have a farming plot of fewer than 2 hectares.

different cities. The three waves of Hukou liberalization, as well as the roll-out of a non-Hukou temporary residence system we consider in this paper, provide significant spatial and temporal variation in policy-based migration costs for farm labor.

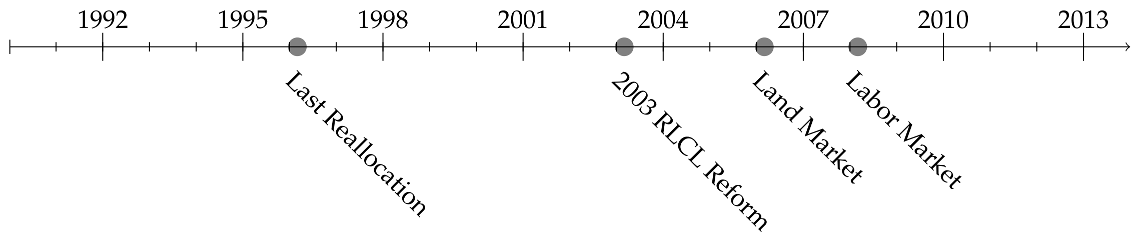
The first wave, the 'Blue Stamp Hukou', was rolled out between 1984 and 1998 and allowed entrepreneurs who made significant investments, white-collar workers, and farmers who had been displaced by government purchases of their land to acquire urban Hukou. Between 1997 and 2001, the second wave allowed migrants who permanently resided in selected (mostly smaller) cities to apply for local Hukou. The third wave extended these regulations to 123 larger cities from 2002 to 2014. In parallel to relaxing Hukou restrictions, some cities introduced temporary residence permits between 1984 and 2000. These allow rural workers to legally reside in the cities but bar them from access to many public services available to urban residents, including schooling, health insurance, social security, and the right to purchase a house. Between 2004 and 2010, some cities strengthened the permit system to allow for permanent residence (Chow, 2015). These staggered timings of reforms at the prefecture level provide exogenous shocks to the cost of migration from rural to urban areas, quasi-randomly change farm laborers' occupational choices, and increase market integration.

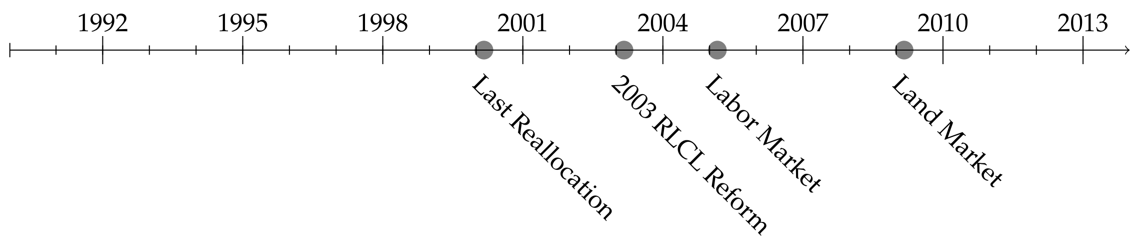
## Land Rental Market Formation

Alongside the missing labor market in rural China, the land market was also incomplete before 2003. The rental activities were based on an oral agreement and shared between family, relatives, and neighbors. The Rural Land Contracting Law (RLCL) implementation in 2003 provided farmers the legal right to rent out and rent in the land, fostering formal contracts and legal security for both parties. Regarding the RLCL, Chari et al. (2021) show that "stated in Article 64, The standing committees of the people's congresses of the provinces, autonomous regions and municipalities directly under the Central Government may, in accordance with this Law and light of the actual conditions of their administrative areas, work out measures for the implementation of this Law." Under these general guidelines, local governments had started rural land contracting reform in a staggered timing, which allows the possibility of identifying the effects of land rental market formation on eliminating or exacerbating the misallocations and distortions.

## When Land Redistribution Meets Market Formation

The following figure shows two villages that experienced the event of the last reallocation, the 2003 RLCL reform, land market formation, and labor market formation in chronological order. As the abolition of land redistribution policy took place at different timings among different villages, this paper focuses on the timing of the last redistribution, defined by the last land reallocation before 2003. It was ex-ante unpredictable since the central government carried out the abandonment of reallocation in 2003. For instance, a household that experienced a redistribution in 1998 cannot infer that this was the "last" redistribution because she cannot predict the coming of the 2003 RLCL reform. Suppose households experienced a family size increase before the last reallocation. In that case, they could get an additional piece of land to match their population. On the contrary, households experienced family size increase after the last reallocation, could not get assigned land, and thus had fewer land plots, even with a similar demographic pattern. Likewise, if households experienced a decrease in family size before the last reallocation, then their land was retrieved to match their population, while households that experienced family size reduction after the last reallocation had no change in their area of land plots. Thus, by this rule, the initial land endowment for each household was "permanently" determined at the timing of the last reallocation, when no market existed.




## 3.3 Theoretical Framework

In this section, first, we theoretically and graphically show how heterogeneity in initial land endowment induced by land reallocation affects household labor allocation across sectors when labor and land markets are restricted. Then we illustrate how the relaxation of labor market restrictions changes the household labor allocation.

### Land Reallocation

When there is no formal land market, each household can only cultivate the assigned amount of land based on the total household size. Hence, every household starts with $land_0$. If the labor market is perfect, labor can switch between sectors without cost, especially agricultural and non-agricultural work. The price of agricultural labor input is the general wage, represented by the slope of the blue line in Figure 3.2, comparable to non-agricultural work. To maximize the total agricultural profits, the household chooses $N_0^*$ as the optimal labor input. However, under the restriction of the Hukou system, rural labor cannot migrate to the non-agricultural sector freely. They invest all the labor force $N_0$ into agriculture. The general wage is the shadow price of labor that households cannot achieve. That is why at the baseline, all households in Figure 3.2a and Figure 3.2a over-invest $N_0 - N_0^*$ labor in the land.

For households experiencing an increase in population before the land reallocation, shown in Figure 3.2a, they can receive extra amounts of land for the village, making their total landholding change from $land_0$ to $land_1$. For households experiencing an increase in household size after the land reallocation, though they have the same household size as the previous group, they still only have $land_0$. More land with the all labor force in agriculture makes the households close to the new optimal labor input $N_1^*$ if the labor market is perfect. Though the households still over-invest labor, the over-supply amount decreases to $N_0 - N_1^*$.

For households experiencing a decrease in the population before the land reallocation, shown in Figure 3.2b, some amounts of land were taken away by the village, making their total landholding change from $land_0$ to $land_2$. With less land, the new optimal labor input is $N_2^*$ if the labor market is perfect. The situation of over-supply of agriculture is magnified if households hold the original agricultural labor input $N_0$ constant, which makes the extra amount of over-supplied labor with even negative marginal utility compared to the leisure time. Therefore, households with less land reduce the agricultural labor input from $N_0$ to $N_2$ to shrink the gap between real labor input and optimal labor input from $N_0 - N_2^*$ to $N_2 - N_2^*$.

If the oversupply of labor in the land is a common phenomenon, in the empirical part, we can see the consistent results with the predictions from the above figures. Households with more land keep the same total labor input as the households have more people after the reallocation. Households with less land reduce their total agricultural labor input.

## Relaxation of Labor Market Restriction

If the labor market restrictions are relaxed, the over-invested labor force in the land can be allocated to non-agricultural sectors and earn the general wage. Shown in Figure 3.3, the previous shadow price now is the real labor price, and the optimal agricultural labor input, such as $N_0^*$, $N_1^*$ in Figure 3.3a, and $N_0^*$, $N_2^*$ in Figure 3.3b can be achieved.

For households with more people without more land in Figure 3.3a, and two groups in the household size decrease group in Figure 3.3b, they all allocated the large oversupplied labor to other sectors. The total labor input in the land decreases by $N_0 - N_0^*$, $N_0 - N_0^*$, and $N_2 - N_2^*$ respectively. For households with more land, since the extra amount of land makes them closer to the optimal point, the agricultural labor input decreases but with less change, $N_0 - N_1^*$.

In the empirical part, after the relaxation of the Hukou system, consistent with the prediction of graphs, we can observe that all groups experience the allocation of the labor force from the agricultural sector to other sectors. Before the reform, the household with more land has the smallest readjustment of the labor force within the four groups.

## 3.4 Data and Identification Strategy

In this section, we describe our main datasets and novel identification strategy, which allow us to identify: 1) effects of initial land endowments on land and labor allocations without a market; 2) effects of land market formation on land and labor allocations; 3) effects of labor market formation on land and labor allocations.

### National Fixed Point Survey (NFPS)

The main data we use is the National Fixed Point Survey (NPFS), which is a nationally representative panel dataset (unbalanced) of 23,000 households in 360 villages between 1986 and 2013. It is collected by the Ministry of Agriculture and Rural Affairs of China, covering all the continental Chinese provinces every year except in 1992 and 1994. It includes 1) a panel of village surveys covering the full sample that provide information on total/detailed agricultural output, consumption, labor demand and input, land lease and use, income, expense, etc; 2) a panel of household surveys consists of 23,000 households that provide information on landholding, input, output consumption, etc; 3) a panel of individual surveys that provide information on born year, gender, educational attainment, disability status, etc. of each individual within a household.

Table 3.1 Panel A presents the summary statistics for the main outcome variables for the full sample of the household survey (N = 36,890). The average household size in the data is 4.13, and the average cultivated land size is 7.86 mu, which is equivalent to 0.52 hectares. Household survey records the land rent activities, including both rent-in and rent-out. More detailed questions about rental activities are only available after 2009. Agricultural income is the main source of total household income. Operational income mainly consists of agricultural income, accounting for 61.3% of the total household income. Questions about labor allocation among sectors are also asked. We select several variables to measure the labor supply decisions. Each household member was asked the number of days they worked off-farm (not in agriculture) in the same township, as a migrant worker, and on-farm (in agriculture). We then aggregate the individual variables to the household level. On average, there are 0.45 people per household working off-farm, earning around 20% of total household income each year. Each household input 202 labor days into agricultural production and 128 labor days into non-agricultural related production each year.

Table 3.1 Panel B shows the summary statistics of the land reallocation scheme at the village level (N = 170). Villages have discretion over the schedule of land reallocation. The decisions were made in the village level representative meeting based on the population change, land resources availability, and aggregated labor supply. There is a lot of variation among villages. Villages experienced 2.4 times land reallocation on average, but some of them only had once reallocation early in 1980, and others had multiple reallocations till 2003 when the national RLCL reform came in. The year gap between two nearby reallocations is four years. The summary statistics of the land reallocation scheme guide our empirical strategy.

## The Timing of the Redistribution and Multiple Reforms

We collect information on the local prefecture-level implementation of the first wave of Hukou reform that was passed at the national level from several different sources. We first collect all migration-related policies using a two-tiered keyword search across multiple platforms to build our policy dataset. In combination with the name / administrative ID of each administrative unit (provinces and prefectures), the keywords we use are:
1. First tier: Rural-urban, urbanized population, Hukou, household registration, living permit, temporary residence, settled down.
2. Second tier: Rural-urban population, Rural-urban residents, migrant workers, migration, registration management, registration reform, registration change, points-based the application system for household registration, abolition of rural and agricultural Hukou.

The primary sources of government regulation documents are from Beidafabao (PKlaw),[7] various governance discussion papers and official government websites, government gazettes, repositories of laws and regulations, as well as documents provided by relevant administrative units. We also complement these sources by directly searching above the keywords through the search engine and historical news. For each document, we digitize 1) when, 2) for whom, and 3) migrating from where.

Our data on the province-level timing of the implementation of the RLCL reform builds on work by Chari et al. (2021), which provides variations in land market frictions. Our data on each round of village-level land redistribution is from the Village Democracy Survey (VDS), built on work by Martinez-Bravo et al. (2017), which records village administrative data and includes information about the timing of each round of reallocation for 183 villages out of 360 villages in the NFPS dataset.

## Identification

As aforementioned, we construct two groups to explore the heterogeneous impact of land endowment differences introduced by the relative exogenous village-level land reallocation on household agricultural production and labor allocation. One is experiencing a household size increase, and the other is experiencing a household size decrease. In each group, there are two subgroups. One is the households with household size change before the land reallocation. Therefore their total land is adjusted during the land reallocation. We name them as a matched group. The other is that the households have household size change after the reallocation. Even though they have similar household structure dynamics as the households with change before the reallocation, a bit late household size change exempts them from adjusting land size during the reallocation. We name them an unmatched group.

Specifically, The treated $T_1$ is defined as groups that had household size increased before the last reallocation. Farmers in $T_1$ were assigned more land to match their family size, while the control $C_1$ corresponds to households that experienced a family size increase after the last reallocation. Households in $C_1$ had fewer land plots than that in

[7]The website is www.pkulaw.com.

$T_1$ due to the redistribution rule, though they had similar demographic dynamics. In parallel, the second treated $T_2$ is defined as groups that decreased household size before the last reallocation, while the second control $C_2$ corresponds to households that experienced family size decrease after the last reallocation. Thus, the initial land endowment for households in $T_2$ was lower than that of $C_2$.

The definitions of these two by two groups are shown in Figure 3.4. The x-axis is the event time relative to the reallocation. Time 0 was the year when the last round of reallocation happened. To avoid overlapping with the previous wave of land reallocation, we cut the pre-reallocation study window at -3, which is three years before the last round of reallocation. The after allocation window, we keep it till 4 to avoid the overlap of afterward land or labor market reform. Both subgroups experience household size increase within the study window for the household size increase. While the matched subgroup experiences a size increase in the -2 to 0 period, the unmatched subgroup experiences a size increase in the 1 to 3 period. Both subgroups experience a household size decrease within the study window for the household size decrease group. While the matched subgroup experiences a size decrease in the -2 to 0 period, the unmatched subgroup experiences a size decrease in the 1 to 3 period. We avoid the 0-1 period to avoid the possibility of a mismatch between the land reallocation timeline and household size change timeline since we do not have the detailed time information of household size change.

## Compliance and Balance Test

Figure 3.5 shows the compliance of land size change in response to family size change before and after the reallocation. We draw the cumulative distribution of households that experienced changes in land plots. The upper curve shows that 85% of households had a decline in land size if their family size decreased before the village-level reallocation. Likewise, more than 70% of households increased land size if their family size increased before the village-level reallocation, as presented in bottom curve.

We then test the balance of the four subgroups ($T_1$ vs. $C_1$ and $T_2$ vs. $C_2$) from household structure, land composition, agricultural and other sector production, labor allocation, expenditure, revenue, etc. at period $t = -2$, which represents two years prior to the last reallocation. Subgroups within each group do not show statistically significant differences. Balance tables are shown in Table C.1 and Table C.2. At the moment of the last reallocation, initial land endowments were determined forever since there is no land-transaction market in China.

We test the balance of the treatment group and control group within each subgroup, increase or decrease, through the family population structures, land resources, agricultural production input and output, family expenditure and income, and labor allocation. Table C.1 shows the comparison within the increased subgroup at the baseline. Table C.2 shows the comparison within the decreased subgroup at the baseline. Since the two groups were in the different stages of family dynamics, household sizes in the increasing group are around 3.9 while household sizes in the decrease group are around 4.5.

However, treatment and control are statistically similar within each group. All groups had around eight mu of land on average, which is close to the level of the full sample. Agricultural income accounts for more than 70% of total household income at the baseline. Compared to the full sample across all periods we studied, the proportion is larger since, with the reforms in labor and land markets, more and more people were working outside agriculture. A similar explanation is with labor allocation across sectors, labor days input in agriculture, and off-farm labor input. The differences between the baseline and full samples prove that families rely less on agriculture.

## 3.5 Empirical Results

In this section, we examine the effects of land redistribution, land market contracts, and labor market formation on aggregate productivity, land size, and labor allocation.

### Redistribution and TFP

We start with a TFP analysis on the effects of land reallocation. The target of the frequent land redistribution is known as to help maintain the egalitarian distribution of land in response to household-level demographic change (Kung, 1994). Then a natural question is whether the redistributions increased the aggregate TFP? We first measure household level TFP using the following equations:

$$y_{it} = \beta_l l_{it} + \beta_n n_{it} + \beta_k k_{it} + \phi_{it}$$

$$\phi_{it} = \alpha_i + \gamma_t + \epsilon_{it}$$

Where $y_{it}$ is the total output of farmer $i$ at time $t$. $l_{it}$, $n_{it}$, and $k_{it}$ represent farmer $i$'s labor input, land, and capital investment, respectively. $\phi_{it}$ is the residual from the first equation, and we decompose it into household fixed effects $\alpha_i$, time fixed effects $\gamma_t$, and the error term $\epsilon_{it}$.

Village-level TFP is measured as the weighted average of household-level TFP using weight $w_{it}$, denoted the share of land plots from household $i$ to the village total land plots.

$$I_{vt} = \sum w_{it}\phi_{it} = E(\phi_{it}|v,t) + \sum (w_{it} - E(w_{it}|v,t))(\phi_{it} - E(\phi_{it}|v,t))$$

Empirically, we estimate the reallocation effects using equations as follows,

$$I_{vt} = \alpha + \beta_1 Reallocation_{vt} + \beta_2 After\ Reallocation_{vt} + \delta_t + \gamma_v + \epsilon_{vt} \tag{3.1}$$

Where $I_{vt}$ represents the weighted aggregate TFP in village $v$ a time $t$. $Reallocation_{vt}$ is an indicator reflecting the year that village reallocated the land plots. $After\ Reallocation_{vt}$ denotes the post-reallocation period for village $v$. $\delta_t$ indicates the time-fixed effects that capture the common trend of each village, and $\gamma_v$ is the village-fixed effects, which control for the time-invariant village-specific characteristics. $\beta_1$ and $\beta_2$ are our coefficients of interest, which could be interpreted as the effects of land redistribution on aggregate TFP.

Table 3.2 reports the regression results of equation (3.1). This evidence shows that the land reallocation did not significantly influence the aggregate TFP in various kinds of measures. The only exception is in Column (1), which suggests that in the year of reallocation, the aggregate TFP was slightly increased by 4 percent, while in the years after the reallocation, TFP had a 6 percent increase. Both of these magnitudes are significant at 10% level. And we do not find similar effects in the other two TFP measures. Overall, we find that redistribution didn't increase productivity instantly or in a longer time period, which again asserts that the goal of redistribution is to achieve egalitarianism instead of efficiency regarding productivity.

## Redistribution and Individual Farm Size

We then test how did redistribution induce heterogeneous treatment effects to land size based on our sample construction described in Section 3.4, which should also suggest the validity of our design.

$$y_{ivt} = \alpha + \beta_3 Pop\ change\ before_{iv} * After\ Reallocation_t + \delta_t + \gamma_i + \epsilon_{ivt} \tag{3.2}$$

Where $y_{ivt}$ is our outcome of interest, like the total size of farmer $i$'s land plots in village $v$ at time $t$. $Pop\ change\ before_{iv}$ is an indicator for the two subgroups: 1) $T_1$ represents households that experienced family size increase before the last reallocation; while 2) $T_2$ indexes households that experienced family size decline before the last reallocation. $Pop\ change\ before_{iv} * After\ Reallocation_t$ is an indicator for the interaction between whether the time is after the last round of land reallocation and the household experiencing pop change before reallocation. $\delta_t$ and $\gamma_i$, agian, are the time fixed effects and farming household fixed effects. $\beta_3$ is the coefficient of interest, which captures the land size change of the households that experienced family size change before the last reallocation compared to that of the households that experienced family size change after the last reallocation. We run separate regressions for the household size increase groups ($T_1$ and $C_1$) and decrease groups ($T_2$ and $C_2$).

Table 3.3 presents the evidence of land size change in the face of the last reallocation. In Columns (1) and (3), the outcome is farm size per capita at the household level, while in Columns (2) and (4), the outcome is farm size per adult labor force. Column (1) shows that, compared to households that had family size increase after the last reallocation ($C_1$), households that experienced family size increase before the last reallocation ($T_1$), on average, were allocated 0.2 mu more land per capita, corresponding to a 10% more land per capita. In terms of land per labor force, $T_1$ got 0.38 mu more per labor compared to that of $C_1$, corresponding to a 12% more land per labor. Symmetrically, for the family size decrease groups, the point estimates in Column (3) and (4) indicate that households that experienced family size decrease before the last reallocation ($T_2$) were allocated 10% less land per capita (0.209/2.091), and 9.5% less land per labor force (0.322/3.402). All of these point estimates are statistically significant at 1% level, suggesting that the exogenous demographic dynamics around the timing of the last reallocation indeed cause a 10% -12% difference in the area of land plots per capita permanently if no market exists.

Then a natural question is how did farmers allocate labor input in response to different land size changes? Panel A of Table 3.4 presents the effects of land reallocation on-farm labor inputs measured by working days, based on equation (3.2), with Columns (1) and (2) for the increase-group comparison and Columns (3) and (4) for the decrease-group comparison. The outcome in Column (1) is the total labor input for the comparison between $T_1$ and $C_1$, while the outcome in Column (2) is the labor input per unit of land. Consistent with our theoretical framework in Section 3.3, the increase in land size didn't cause $T_1$ to increase total labor supply (Column 1), since they already oversupplied labor in agriculture , but it indeed reduced the labor input per unit of land by 7.6% because $T_1$ received more land compared to $C_1$. Similarly, the outcome in Column (3) is the total labor input for comparing $T_2$ and $C_2$, while the outcome in Column (4) is the labor input per unit of land. In contrast, consistent with our model, farmers in $T_2$ reduced total labor input in agriculture by 14.5% relative to $C_2$ because of losing land (Column 3) , while their labor input per unit of land shows no different, compared to $C_2$, because they simultaneously lost land and family size (matched group).

Panel B of Table 3.4 presents the effects of land reallocation on the non-agriculture job within the village, based on equation (3.2). Columns (1) and (2) show the increase-group comparison, and Columns (3) and (4) present the decrease-group comparison. In Column(1), we do find a 21.5% less allocation in non-agricultural labor input for the $T_1$-$C_1$ comparison, which is statistically significant at 5% level. This difference is driven by the farmers in $C_1$ group, since they had a larger family size without being allocated more land. To feed those extra population, they need to input more labor into the non-agriculture sector within the village since rural-urban migration in the same period was extremely restricted.

## Land Market Formation

We then conduct an analysis on how did land market reform reshape household land size and labor input using the specification (3.3) below,

$$\begin{aligned} y_{ivt} = & \alpha + \beta_4 Reform_{vt} * Pop\ change\ before_{iv} + \\ & \beta_5 Reform_{vt} * Pop\ change\ after_{iv} + \delta_t + \gamma_i + \epsilon_{ivt} \end{aligned} \tag{3.3}$$

where $y_{ivt}$ is the outcome of interest, including land size per capita, land size per labor, and agricultural labor input. $Reform_{vt}$ is a dummy indicator that represents the staggered timing of the implementation of RLCL reforms, which takes the value of 1 if the villages were at the post-reform period. $Reform_{vt} * Pop\ change\ before_{iv}$ represents the interaction between the timing of land reform and two treatment subgroups, including 1) $T_1$ representing households that experienced family size increase before the last reallocation and 2) $T_2$ indexes households that experienced family size decline before the last reallocation. $Reform_{vt} * Pop\ change\ after_{iv}$ represents the interaction between the timing of land reform and two control subgroups, including 1) $C_1$ representing households that experienced family size increase after the last reallocation and 2) $C_2$ indexing households that experienced family size decline after the last reallocation. $\delta_t$ and $\gamma_i$ are the time fixed effects and farming household fixed effects. $\beta_4$ and $\beta_5$ are the coefficients of interest, which capture the heterogeneous impacts of land reform on households with different initial land endowments.

Table 3.5 presents the point estimates of the impact of land reform on household land size using equation (3.3). Columns (1) and (2) show the results of the size increase group, while Columns (3) and (4) demonstrate the results of the size decrease group. Column (1) suggests that, compared to the pre-reform period, the formation of the land market increases the land plots per capita of households with a low initial land endowment (C1) by 5%, and land per labor by 7.4%. Intuitively, such a rise in land plots results from rent-in activities, since in $C_1$ groups, households experienced population size increase but were not assigned additional land plots. They should have more demand for rent-in land for agriculture to feed more people. In contrast, we do not see an increase in land per capita for households in T1 group after the land contract reform, which experienced simultaneous population size increase and land size increase. Farmers in this group showed less interest to rent in more land since their family size and land size were already matched.

Turn to Columns (3) and (4) in Table 3.5, we find that compared to the pre-reform period, the formation of the land market doesn't increase the land plots per capita/labor for households that experienced family size decrease before/after the last reallocation. For the households that experienced a family size decrease before the last reallocation ($T_2$), their land was also retrieved accordingly, which balances the population size and land size. Thus they had no incentive to rent the land after the land reform. For the households that experienced a family size decrease after the last reallocation ($C_2$), their land per capita/labor was even higher than that of ($T_2$), causing low demand for land rental activities.

In table 3.6, we examine how land reform reshapes household factor input allocation. In Columns (1) and (2), the outcomes are labor input and other input for the size increase group. We do not find evidence that land reform changes household factor inputs. Similarly, Columns (3) and (4) also show that households that experienced a family size decrease before/after the last reallocation didn't adjust their labor input and other input according to land reform reasons. This evidence, consistent with our theoretical framework, suggests that these households had already oversupplied labor in agriculture. They do not need to increase labor input when they can rent more land after the land reform.

Overall, we find that the formation of land and market contracts only affects households with low initial land endowment but increasing family size, allowing them to rent land to maintain the consumption of a larger family size. But we hardly see an impact of land reform on agricultural labor supply due to pre-existing over-employment of agricultural labor.

### Labor Market Formation

We also analyze the effects of labor market formation on household labor allocation, using the same equation (3.3). Again $\beta_4$ and $\beta_5$ are our coefficients of interest, which capture the differences in the labor outcomes between the pre-reform period and post-reform period. Table 3.7 shows the point estimates of the effects of Hukou reform on household labor allocation in agriculture. Columns (1) and (2) show the results of the size increase group, while Columns (3) and (4) demonstrate the results of the size decrease group. In Columns (1) and (2), the point estimates are -0.73 and -0.47 (at 5% level), indicating that Hukou reform has significantly led to a lower agricultural labor supply for the households that had family size increase after the last land reallocation. And we do not find a significant change in agricultural labor supply for the households that had family size increase before the last land reallocation since their land size and family size were already matched. Similarly, in Columns (3) and (4), labor supply in agriculture dropped by roughly 70% in total and 50% per unit of land for the households that experienced a population size decrease, no matter before or after the last land reallocation.

Last, we estimate the interacting effects between Hukou reform and land contracting reform by pooling two reforms in the regression. Our goal is to help policymakers to decide which markets to prioritize. In Table 3.8, Columns (1) and (2) report the pooled effects for the increasing group, while Columns (3) and (4) present the pooled effects for the decreasing group. This point estimates consistently suggest that land reform seems less effective in reshaping household labor allocation in the presence of labor market reform. For both increase and decrease groups, Hukou reform consistently reduces total agricultural labor supply by roughly 45 - 54% causing more migrant workers to work in the urban area and inducing structural transformation. A considerable distortion behind such change is that households oversupplied labor in the agricultural sector and acquired low marginal product of labor. And the land reform can increase TFP and the marginal product of labor via increasing the size of land per capita, while the labor market reform increases TFP through reallocation of labor from agriculture to urban migrant work.

## 3.6 Robustness Check

This section examines the robustness of our results to alternative specifications and explanations. First, we consider the possibility that households might have strategic behavior before the last land reallocation. If households form expectations on the reallocation, they might adjust their family size before the reallocation. In fact, we just focus on the last round of land reallocation. It is unlikely that households in the year before 2003 could predict that that was the last reallocation for two reasons. First, the land reform is a top-down reform initiated by the central government in 2003, prohibiting land reallocation nationwide. The provincial leaders then follow a centralized plan, and city leaders follow the province strategy. It is almost impossible for a household in the village to foresee the incoming land contracting reform and think that it was the last round. As shown in Figure

3.1, the frequency of land reallocation dropped to three substantially after 2003.

Second, Based on the pre-determined land reallocation rule, it is difficult for a household to manipulate the reallocation decision since it requires the involvement of all the households in the village. As shown in Zhao (2020), the exact procedure by which land is redistributed across households varies across villages. However, a common practice is for village leaders to redivide all land in the village with equal distribution of land quality types and then distribute these bundles to all households based on a pre-determined allocation rule. Figure C.1 shows that the shares of households that experienced family size increase, decrease, and no change. We find no evidence that households demonstrated strategic behavior around the timing of the last land reallocation. In figure C.2, we present the village-average demographic dynamics of households in each village near the timing of land reallocation, which is quite smooth and continuous.

Besides, our identification could suffer from the threat of the definition of treatment group ($T_1, T_2$), control group ($C_1, C_2$), and the reallocation time window. To rule out the possibility that the results are driven by the specification of the study window and group definition, we do a robustness check by including $T - 4$ or $T + 5$ for the period under restrictive markets and also using the more stringent definitions of groups for both the restrictive period and post-reform period.

We extend the study window to $T-4$ or $T+5$ separately and estimate the effects of land reallocation on the land endowment and labor allocation for the period when both labor and land market are imperfect. The results of land size change are shown in Table C.4 in Appendix C. The magnitude and statistical significance are consistent with our main results. We then test the household labor allocation.

We then use a more stringent definition of control and treatment group by limiting the samples to households that experienced population change only just before one year and after one year around the land reallocation for this robustness practice. We test both the effects of land reallocation and two market reforms. The results of the effects of land reallocation on land size change and labor allocation are presented in Table C.5. The results of the effects of land reform on land size are presented in Table C.6 and results of the effects of Hukou system reform on labor allocation are presented in Table C.7.

## 3.7 Conclusion

Developing countries are characterized by substantially heterogeneous farm sizes across individual farmers. There is a long debate that small-size farmers present higher productivity than large landholders. One of the potential explanations is that labor-market transaction costs induce the slightly larger farms to be least efficient. Why are there productivity differences across different farm sizes? Identifying the causal relationship between land endowment and households labor allocation is notoriously challenging since it requires: 1) a setting or randomized controlled trial that endogenously assigns heterogeneous farm sizes to different households; 2) a system where both labor and land markets are malfunctioning; 3) a sufficiently long track of many individual households.

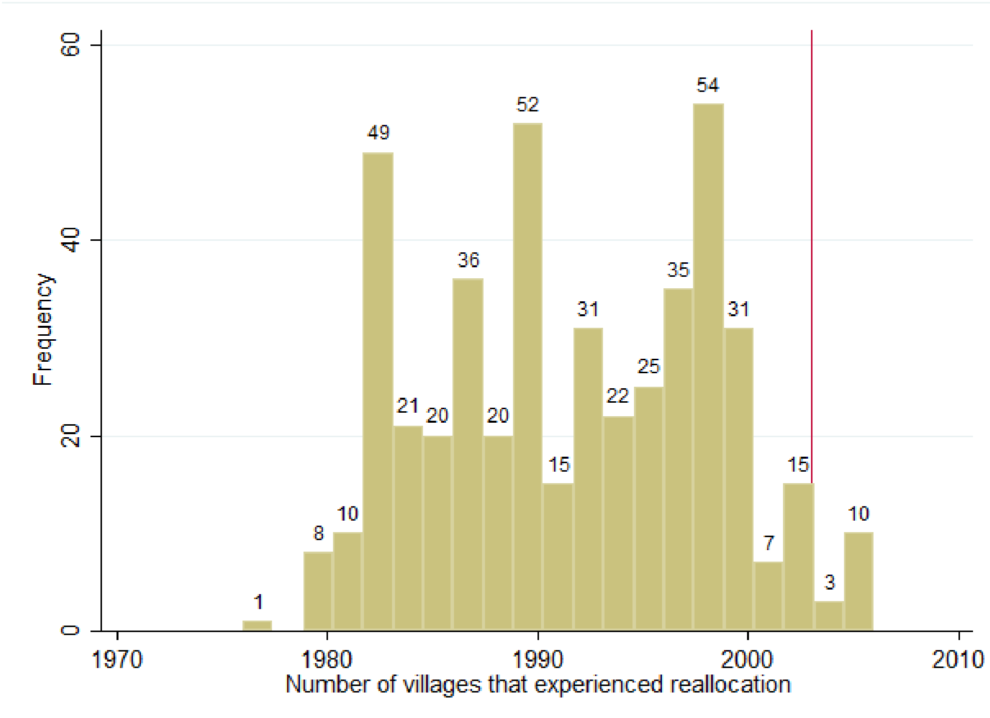
Small literature by Gottlieb and Grobovšek (2019); Foster and Rosenzweig (2022) has begun to explore the role of farm size in productivity growth. But we still know little about the effect of an exogenous land endowment on household long-run performance and welfare, the underlying mechanisms, and policies that can correct such a change.

Using a unique land redistribution setting in China, this paper investigates the effects of initial land endowment on household labor allocation when no land and labor market exists. It further evaluates the impact of land contracting law reform and labor market reform on labor allocation decisions. We first find that in the absence of the land market and labor market, the size of households before the last land redistribution determined the land size permanently, causing households to oversupply labor in agriculture. Consequently, the land reallocation scheme did not increase the village-level aggregate TFP.

We then find heterogeneous treatment effects of land contracting reform on household land size since the formation of the land market allowed households to rent in and rent out land formally and legally. Unlike previous literature that farmers make better decisions on labor input when the land market is complete, we find insignificant and small effects of land reform on labor allocation. The driver of such a puzzle is consistent with our statement that farmers in China were oversupplying labor in agriculture. As a result, the marginal product of labor was lower than the optimal value. Land market reform increased overall TFP by reallocating land to those with large family sizes but low initial land endowment. Last, we also examine the impact of rural-urban labor market reform on household labor allocation. Evidence shows that reducing the mobility restrictions of rural-urban migration significantly reduced labor supply in agriculture and led to a higher marginal product of labor in agriculture and structural transformation.

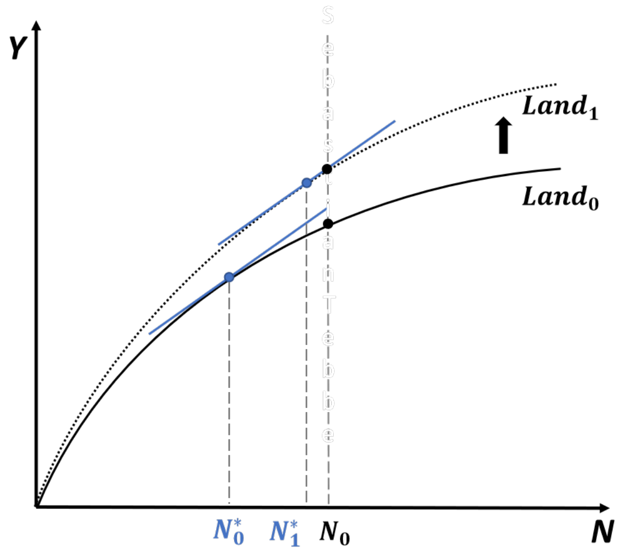
As for policy-makers, a central question is which reform should be prioritized when there are simultaneous missing land and labor markets. By introducing the interaction between the two reforms— land market reform and labor market reform, we find that labor market reform plays a leading role in reallocating labor inputs in agriculture and urban work relative to land market reform. Of course, the key question for a complete evaluation of China's land market and labor is the extent to which the land redistribution is inefficient and what are the welfare implications of the two market reforms. How do these two types of reforms affect households' decision-making in investment and consumption? These are important and interesting questions for future research.

Figure 3.1: Number of Villages that Experienced Redistribution Across Years: N=183

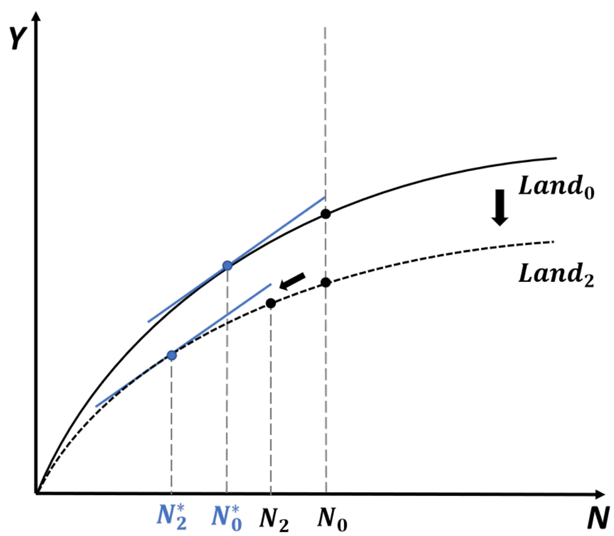



Note: This figure presents the number of villages that experienced redistribution in 183 villages in each year. The red line indicates the timing of the implementation of 2003 Rural Land Contracting Law.

Figure 3.2: Land Reallocation under Land and Labor Market Restrictions

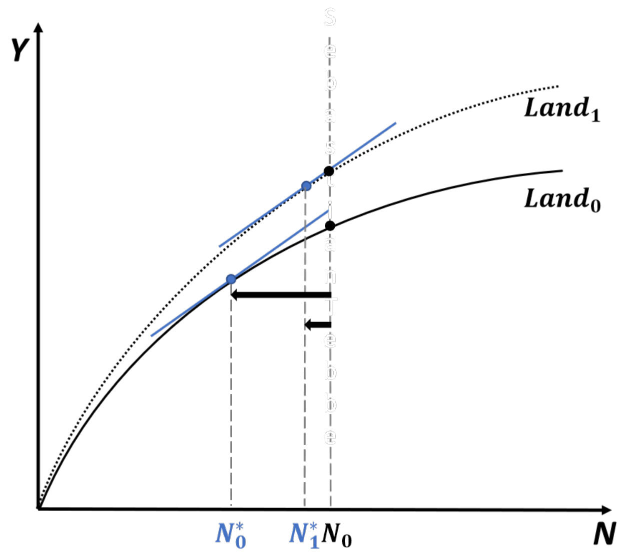



(a) Household Size Increase Group

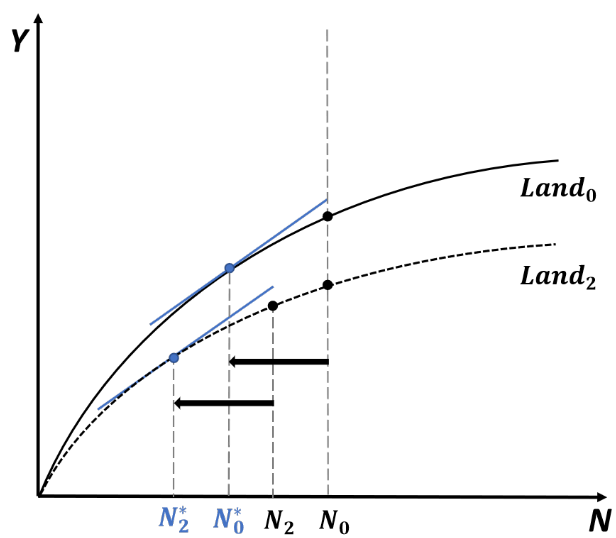



(b) Household Size Decrease Group

Figure 3.3: Land Reallocation under Land Market Restrictions (Labor Restriction Relaxed)

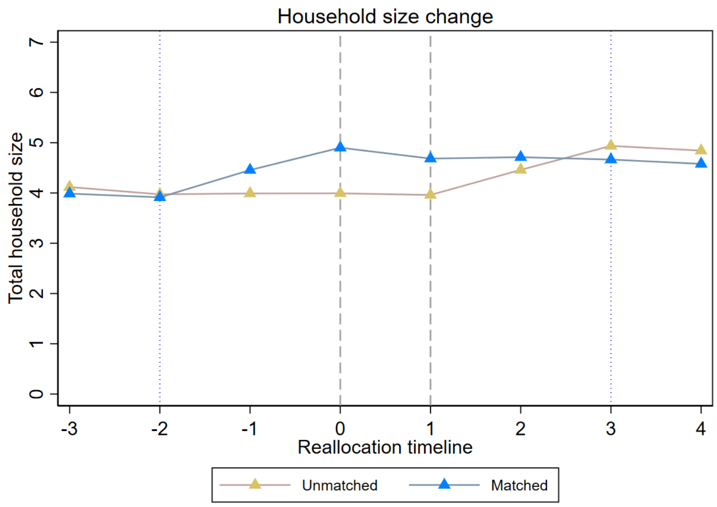



(a) Household Size Increase Group

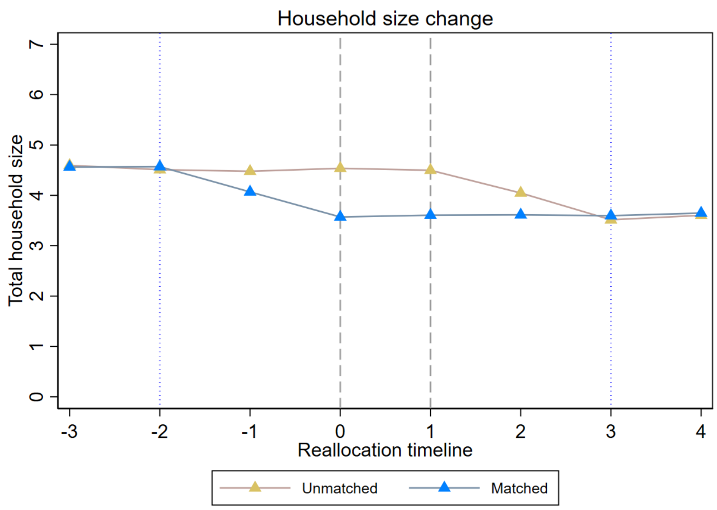



(b) Household Size Decrease Group

Figure 3.4: The Definitions of Four Groups

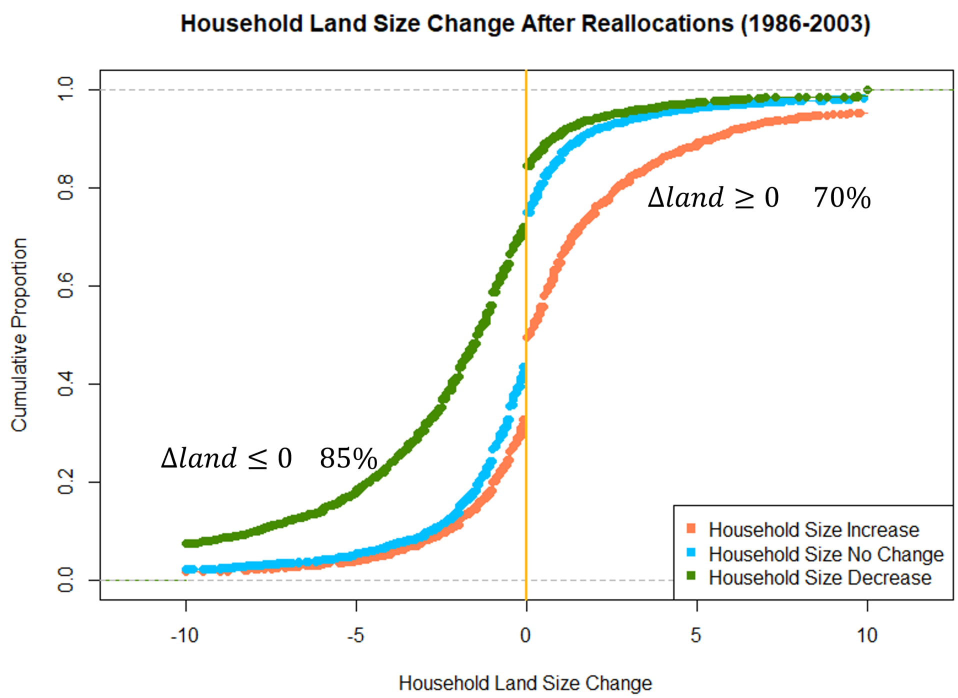



(a) Household Size Increase Group




(b) Household Size Decrease Group

Figure 3.5: Land Size Change in Conjunction with Family Size Change




Note: This figure shows that how land size responds to family size change before and after the reallocation. The y-axis indexes the cumulative distribution and x-axis represents the change in household land size, divided by zero line. We could observe that 85% of households had a decline in land size if their family size decreased before the village-level reallocation (blue). Likewise, more than 70% of households had a increase in land size if their family size increased before the village-level reallocation (orange).

Table 3.1: Summary Statistics

| | (1) Mean | (2) SD | (3) Min | (4) Max | (5) N |
|---|---|---|---|---|---|
| **Household Variables** | | | | | |
| Household Size | 4.13 | 1.55 | 0.00 | 53.00 | 36,890 |
| Total land size (mu) | 7.86 | 11.05 | 0.00 | 306.00 | 36,879 |
| Pieces of land | 5.56 | 6.23 | 0.00 | 99.00 | 36,890 |
| Rent-in land size (mu) | 0.48 | 2.34 | 0.00 | 75.00 | 18,960 |
| Rent-out land size (mu) | 1.20 | 39.94 | 0.00 | 5,500.00 | 19,567 |
| Total Income | 24,598.75 | 44,406.16 | 0.00 | 4,505,100.00 | 36,884 |
| Total operation income | 15,078.40 | 39,985.09 | 0.00 | 4,500,000.00 | 36,890 |
| Labor working off-farm | 0.45 | 0.79 | 0.00 | 8.00 | 36,884 |
| Migrating labor income | 4,935.66 | 10,532.23 | 0.00 | 300,000.00 | 36,890 |
| Total labor input (days) | 443.35 | 5,423.82 | 0.00 | 1,024,225.00 | 36,890 |
| Ag labor input (days) | 202.21 | 237.54 | 0.00 | 30,151.00 | 36,890 |
| Off-farm labor input (days) | 147.55 | 215.38 | 0.00 | 3,000.00 | 16,934 |
| **Village Variables** | | | | | |
| Year of last reallocation | 1,995.55 | 5.24 | 1,980.00 | 2,003.00 | 170 |
| Number of land reallocation | 2.44 | 1.41 | 1.00 | 8.00 | 170 |
| Gap between reallocations (last) | 7.13 | 4.26 | 2.00 | 18.00 | 121 |
| Gap between reallocations | 4.18 | 1.85 | 1.00 | 9.00 | 121 |

Table 3.2: Village Level Aggregate Revenue TFP

| | (1) | (2) | (3) |
|---|---|---|---|
| | Aggregate TFP | Average TFP | OP Cov |
| Reallocation | 0.042* | 0.040 | 0.002 |
| | (0.025) | (0.031) | (0.018) |
| After Reallocation | 0.062* | 0.036 | 0.026 |
| | (0.037) | (0.040) | (0.029) |
| Household FE | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes |
| Observations | 979 | 979 | 979 |

*Note:* This table presents the village-level TFP in face of land redistribution and the change of village-level TFP after land redistribution. In columns (1), (2) and (3), we add both household fixed effects and year fixed effects. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 3.3: Household Land Size Change in Response to Reallocation

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | HH Increase | | HH Decrease | |
| | Per Capita | Per labor | Per Capita | Per labor |
| Pop change before | 0.193*** | 0.383*** | -0.209*** | -0.322** |
| reallocation | (0.065) | (0.136) | (0.076) | (0.138) |
| Y mean | 1.905 | 3.146 | 2.091 | 3.402 |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 7147 | 7051 | 10336 | 10133 |

*Note:* This table presents the point estimate of household land size in response to each round of land redistribution. The outcome variables in columns (1) and (2) are the land size per capita and land size per labor for the household size increase group, respectively. The outcome variables in columns (3) and (4) are the land size per capita and land size per labor for the household size decrease group, respectively. In columns (1)- (4), we add both household fixed effects and year fixed effects.
Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 3.4: Household Labor Inputs (workdays) in Agriculture vs. Non-agriculture

| | HH Increase | | HH Decrease | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A: in agriculture** | | | | |
| | Total | Per land | Total | Per land |
| Pop change before reallocation | -0.068 | -0.076* | -0.145** | -0.032 |
| | (0.057) | (0.044) | (0.062) | (0.044) |
| Y mean | 429.3686 | 107.286 | 418.622 | 105.4928 |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 7147 | 6901 | 10336 | 9887 |
| **Panel B: non-agricultural job** | | | | |
| | Total Nonag | Outside Village | Total Nonag | Outside Village |
| Pop change before reallocation | -0.283** | 0.033 | -0.082 | 0.010 |
| | (0.126) | (0.094) | (0.106) | (0.106) |
| Y mean | 53.860 | 29.155 | 44.700 | 24.411 |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 6901 | 6901 | 9887 | 9890 |

*Note:* This table presents the difference in household labor input between the constructed treatment and control groups. Panel A reports the point estimates of household labor input in agriculture, while panel B reports the point estimates of household labor input in non-agricultural job within the village. In Panel A, the outcome variables in columns (1) and (2) are the total labor input and labor input per unit of land for the household size increase group, respectively. The outcome variables in columns (3) and (4) are the total labor input and labor input per unit of land for the household size decrease group, respectively. In Panel B, the outcome variables in columns (1) and (2) are the total labor input in non-agricultural job within and outside the village for the household size increase group, respectively. The outcome variables in columns (3) and (4) are the total labor input in non-agricultural job within and outside the village for the household size decrease group, respectively. In columns (1)- (4), we add both household fixed effects and year fixed effects.
Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 3.5: Household Level Land Size Change

| | HH Increase (+) | | HH Decrease (-) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Asinh Per Capita | Asinh Per labor | Asinh Per Capita | Asinh Per labor |
| Reform × Pop change after | 0.050** (0.022) | 0.074*** (0.017) | -0.007 (0.023) | -0.001 (0.022) |
| Reform × Pop change before | 0.001 (0.019) | -0.017 (0.022) | 0.034 (0.025) | 0.042 (0.026) |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 5152 | 4911 | 7252 | 6812 |

*Note:* This table presents the point estimate of household land size in response to land reforms. The outcome variables in columns (1) and (2) are the land size per capita and land size per labor for the household size increase group, respectively. The outcome variables in columns (3) and (4) are the land size per capita and land size per labor for the household size decrease group, respectively. In columns (1)- (4), we add both household fixed effects and year fixed effects. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 3.6: Household Labor (days) Allocation

| | HH Increase (+) | | HH Decrease (-) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Ag labor | Migration Inc | Ag labor | Migration Inc |
| Reform × | 0.014 | -0.017 | 0.132 | -0.199 |
| Pop change after | (0.148) | (0.260) | (0.120) | (0.215) |
| Reform × | 0.012 | -0.084 | 0.126 | -0.015 |
| Pop change before | (0.096) | (0.219) | (0.149) | (0.318) |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 5165 | 5165 | 7253 | 7253 |

*Note:* This table presents the point estimate of household labor input in response to land reforms. The outcome variables in columns (1) and (2) are the labor input in agriculture and migrant income for the household size increase group, respectively. The outcome variables in columns (3) and (4) are the labor input in agriculture and migrant income for the household size decrease group, respectively. In columns (1)- (4), we add both household fixed effects and year fixed effects.
Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 3.7: Household Level Agricultural Labor Input Per Land

| | HH Increase (+) | | HH Decrease (-) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Total | Per land | Total | Per land |
| Reform × | -0.725** | -0.465** | -0.782*** | -0.512*** |
| Pop change after | (0.293) | (0.212) | (0.236) | (0.152) |
| Reform × | -0.170 | -0.116 | -0.721*** | -0.471*** |
| Pop change before | (0.196) | (0.174) | (0.237) | (0.151) |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 3922 | 3922 | 5840 | 5840 |

*Note:* This table presents the point estimate of household labor input in response to labor market reforms. The outcome variables in columns (1) and (2) are the total labor input in agriculture and agricultural labor input per unit of land for the household size increase group, respectively. The outcome variables in columns (3) and (4) are the total labor input in agriculture and agricultural labor input per unit of land for the household size decrease group, respectively. In columns (1)-(4), we add both household fixed effects and year fixed effects.
Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table 3.8: Household Agricultural Labor Input Per Land

| | HH Increase (+) | | HH Decrease (-) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Change before | Change after | Change before | Change after |
| Land Reform | 0.047 | -0.157 | 0.042 | 0.107 |
| | (0.184) | (0.213) | (0.205) | (0.142) |
| Hukou Reform | -0.287 | -0.446* | -0.499*** | -0.538*** |
| | (0.174) | (0.248) | (0.188) | (0.173) |
| Land Reform | 0.119 | 0.263 | 0.033 | 0.079 |
| × Hukou Reform | (0.176) | (0.224) | (0.177) | (0.137) |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 2056 | 1817 | 2918 | 2826 |

*Note:* In the regression, we pool land market reform and labor market reform together. This table presents the point estimate of household labor input in response to labor market reforms. The outcome variables in columns (1) - (4) are the total labor input in agriculture. In columns (1)-(4), we add both household fixed effects and year fixed effects.

Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

# Bibliography

Adamopoulos, Tasso, and Diego Restuccia. 2014. “The size distribution of farms and international productivity differences.” *American Economic Review*, 104(6): 1667–97.

Adamopoulos, Tasso, Loren Brandt, Chaoran Chen, Diego Restuccia, and Xiaoyun Wei. 2022. “Land Security and Mobility Frictions.” National Bureau of Economic Research.

Adamopoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia. 2017. “Misallocation, selection and productivity: A quantitative analysis with panel data from China.” National Bureau of Economic Research.

Agyei-Holmes, Andrew, Niklas Buehren, Markus Goldstein, Robert Darko Osei, Isaac Osei-Akoto, and Christopher Udry. 2020. “The Effects of Land Title Registration on Tenure Security, Investment and the Allocation of Productive Resources.” *Global Poverty Research Lab Working Paper*, , (20-107).

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

Aragón, Fernando M, Diego Restuccia, and Juan Pablo Rud. 2022. “Assessing misallocation in agriculture: plots versus farms.” National Bureau of Economic Research.

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.

Banerjee, Abhijit V, Paul J Gertler, and Maitreesh Ghatak. 2002. “Empowerment and efficiency: Tenancy reform in West Bengal.” *Journal of political economy*, 110(2): 239–280.

Banks, James, Leandro S Carvalho, and Francisco Perez-Arce. 2019. “Education, Decision Making, and Economic Rationality.” *Review of Economics and Statistics*, 101(3): 428–441.

Bauman, Yoram, and Elaina Rose. 2011. “Selection or indoctrination: Why do economics students donate less than the rest?” *Journal of Economic Behavior & Organization*, 79(3): 318–327.

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.

Benjamin, Dwayne. 1992. "Household composition, labor markets, and labor demand: testing for separation in agricultural household models." *Econometrica: Journal of the Econometric Society*, 287–322.

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

Bils, Mark, Peter J Klenow, and Cian Ruane. 2021. "Misallocation or mismeasurement?" *Journal of Monetary Economics*, 124: S39–S56.

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.

Brainerd, Elizabeth, and Nidhiya Menon. 2014. "Seasonal effects of water quality: The hidden costs of the Green Revolution to infant and child health in India." *Journal of Development Economics*, 107: 49–64.

Bryan, Gharad, and Melanie Morten. 2019. "The aggregate productivity effects of internal migration: Evidence from Indonesia." *Journal of Political Economy*, 127(5): 2229–2268.

Camerer, Colin F. 1987. "Do biases in probability judgment matter in markets? Experimental evidence." *The American Economic Review*, 77(5): 981–997.

Cantoni, Davide, Yuyu Chen, David Y Yang, Noam Yuchtman, and Y Jane Zhang. 2017. "Curriculum and ideology." *Journal of Political Economy*, 125(2): 338–392.

Card, David, Stefano DellaVigna, Patricia Funk, and Nagore Iriberri. 2020. "Are referees and editors in economics gender neutral?" *The Quarterly Journal of Economics*, 135(1): 269–327.

Carrell, Scott E, Marianne E Page, and James E West. 2010. "Sex and science: How professor gender perpetuates the gender gap." *The Quarterly Journal of Economics*, 125(3): 1101–1144.

Carter, John R, and Michael D Irons. 1991. "Are economists different, and if so, why?" *The Journal of Economic Perspectives*, 5(2): 171–177.

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*.

Chari, Amalavoyal, Elaine M Liu, Shing-Yi Wang, and Yongxiang Wang. 2021. "Property rights, land misallocation, and agricultural efficiency in China." *The Review of Economic Studies*, 88(4): 1831–1862.

Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis. 2021. "The effects of land markets on resource allocation and agricultural productivity." *Review of Economic Dynamics*.

Chen, M Keith, Venkat Lakshminarayanan, and Laurie R Santos. 2006. "How basic are behavioral biases? Evidence from capuchin monkey trading behavior." *Journal of Political Economy*, 114(3): 517–537.

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.

Chen, Yan, and Onur Kesten. 2017. "Chinese college admissions and school choice reforms: A theoretical analysis." *Journal of Political Economy*, 125(1): 99–139.

Chen, Yan, and Onur Kesten. 2019. "Chinese college admissions and school choice reforms: An experimental study." *Games and Economic Behavior*, 115: 83–100.

Chow, Gregory C. 2015. *China's economic transformation.* John Wiley & Sons.

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.

De Janvry, Alain, Kyle Emerick, Marco Gonzalez-Navarro, and Elisabeth Sadoulet. 2015. "Delinking land rights from land use: Certification and migration in Mexico." *American Economic Review*, 105(10): 3125–49.

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).

Faravelli, Marco. 2007. "How context matters: A survey based experiment on distributive justice." *Journal of Public Economics*, 91(7-8): 1399–1422.

Field, Erica. 2007. "Entitled to work: Urban property rights and labor supply in Peru." *The Quarterly Journal of Economics*, 122(4): 1561–1602.

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.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2009. "Exposure to ideology and distributional preferences." *Unpublished paper*.

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

Foster, Andrew D, and Mark R Rosenzweig. 2022. "Are There Too Many Farms in the World? Labor Market Transaction Costs, Machine Capacities, and Optimal Farm Size." *Journal of Political Economy*, 130(3): 000–000.

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

Gai, Qingen, Naijia Guo, Bingjing Li, Qinghua Shi, Xiaodong Zhu, et al. 2021. "Migration Costs, Sorting, and the Agricultural Productivity Gap."

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

Goldstein, Markus, and Christopher Udry. 2008. "The profits of power: Land rights and agricultural investment in Ghana." *Journal of political Economy*, 116(6): 981–1022.

Gollin, Douglas, and Christopher Udry. 2021. "Heterogeneity, measurement error, and misallocation: Evidence from African agriculture." *Journal of Political Economy*, 129(1): 1–80.

Gottlieb, Charles, and Jan Grobovšek. 2019. "Communal land and agricultural productivity." *Journal of Development Economics*, 138: 135–152.

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*a*. "Twenty-year economic impacts of deworming." *Proceedings of the National Academy of Sciences*, 118(14).

Hamory, Joan, Marieke Kleemans, Nicholas Y Li, and Edward Miguel. 2021*b*. "Reevaluating agricultural productivity gaps with longitudinal microdata." *Journal of the European Economic Association*, 19(3): 1522–1555.

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.

Holt, Charles A, and Susan K Laury. 2002. "Risk aversion and incentive effects." *American economic review*, 92(5): 1644–1655.

Hossain, Tanjim, and Ryo Okui. 2013. "The binarized scoring rule." *Review of Economic Studies*, 80(3): 984–1001.

Imbert, Clement, Marlon Seror, Yifan Zhang, and Yanos Zylberberg. 2018. "Migrants and firms: Evidence from china."

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.

Jacoby, Hanan G, Guo Li, and Scott Rozelle. 2002. "Hazards of expropriation: tenure insecurity and investment in rural China." *American Economic Review*, 92(5): 1420–1447.

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*.

Kagel, John H, Chung Kim, and Donald Moser. 1996. "Fairness in ultimatum games with asymmetric information and asymmetric payoffs." *Games and Economic Behavior*, 13(1): 100–110.

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

Kim, Hyuncheol Bryant, Syngjoo Choi, Booyuel Kim, and Cristian Pop-Eleches. 2018. "The role of education interventions in improving economic rationality." *Science*, 362(6410): 83–86.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad. 2016. "Field of study, earnings, and self-selection." *The Quarterly Journal of Economics*, 131(3): 1057–1111.

Krupka, Erin L, and Roberto A Weber. 2013. "Identifying social norms using coordination games: Why does dictator game sharing vary?" *Journal of the European Economic Association*, 11(3): 495–524.

Kung, James Kaising. 1994. "Egalitarianism, subsistence provision, and work incentives in China's agricultural collectives." *World Development*, 22(2): 175–187.

LaFave, Daniel, and Duncan Thomas. 2016. "Farms, families, and markets: New evidence on completeness of markets in agricultural settings." *Econometrica*, 84(5): 1917–1960.

Lührmann, Melanie, Marta Serra-Garcia, and Joachim Winter. 2018. "The impact of financial education on adolescents' intertemporal choices." *American Economic Journal: Economic Policy*, 10(3): 309–32.

Manysheva, Kristina. 2021. "Land Property Rights, Financial Frictions, and Resource Allocation in Developing Countries." mimeo.

Martinez-Bravo, Monica, Gerard Padró I Miquel, Nancy Qian, and Yang Yao. 2017. "The rise and fall of local elections in China: theory and empirical evidence on the autocrat's trade-off." National Bureau of Economic Research.

Marwell, Gerald, and Ruth E Ames. 1981. "Economists free ride, does anyone else?: Experiments on the provision of public goods, IV." *Journal of public economics*, 15(3): 295–310.

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.

Ngai, L Rachel, Christopher A Pissarides, and Jin Wang. 2019. "China's mobility barriers and employment allocations." *Journal of the European Economic Association*, 17(5): 1617–1653.

Rabin, Matthew. 2000. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." *Econometrica*, 68(5): 1281–1292.

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

Rozelle, Scott, and Guo Li. 1998. "Village leaders and land-rights formation in China." *The American economic review*, 88(2): 433–438.

Sarsons, Heather. 2017. "Recognition for group work: Gender differences in academia." *American Economic Review*, 107(5): 141–45.

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.

Tombe, Trevor, and Xiaodong Zhu. 2019. "Trade, migration, and productivity: A quantitative analysis of china." *American Economic Review*, 109(5): 1843–72.

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.

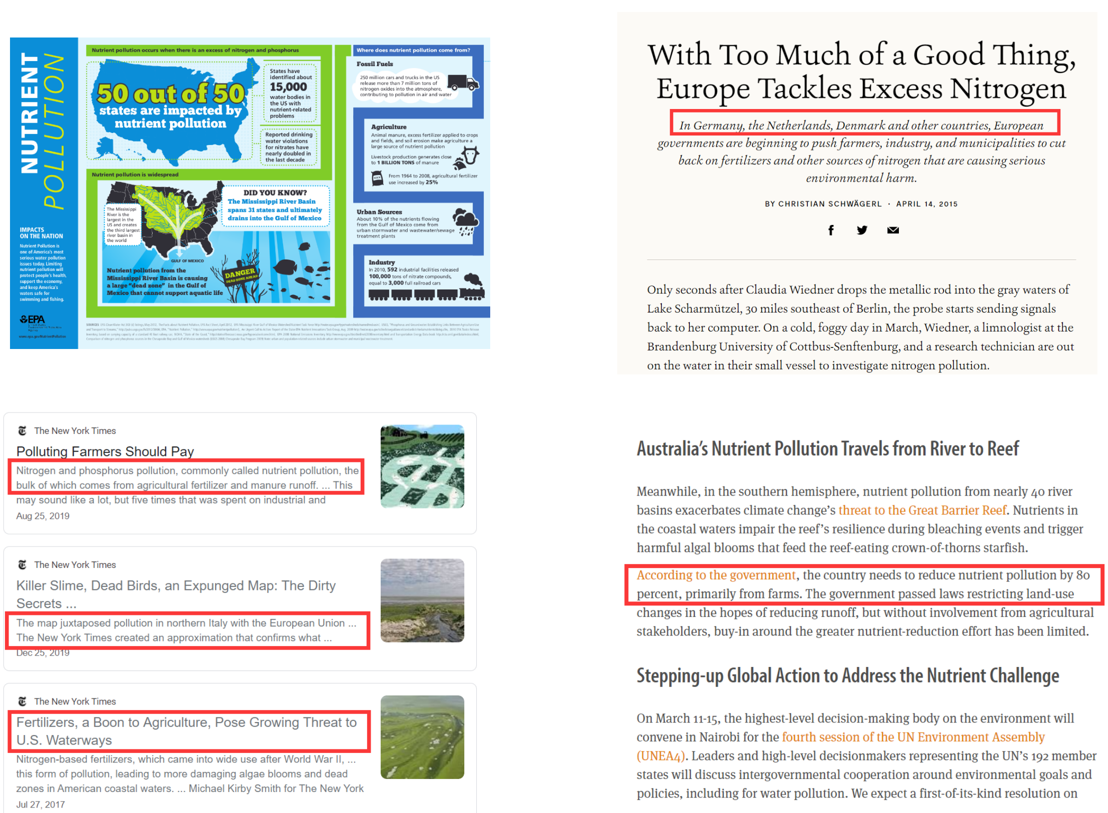
Zhao, Xiaoxue. 2020. "Land and labor allocation under communal tenure: Theory and evidence from China." *Journal of Development Economics*, 147: 102526.

# Appendix A

# Chapter 1 Appendix

## A.1 Supplementary Tables and Figures

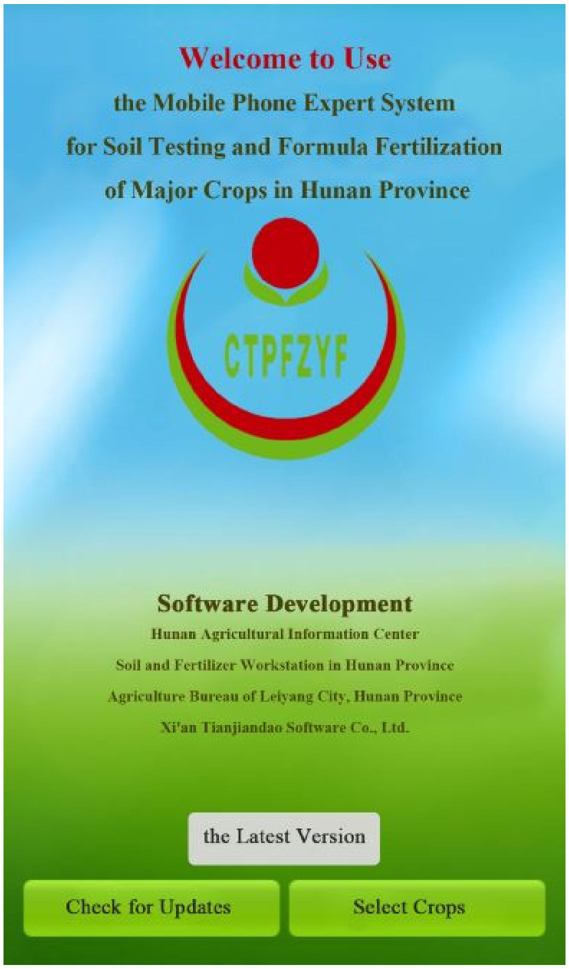
Figure A.1: Nitrogen Abuse in Developed Countries




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

Figure A.2: 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 A.3: The Second Two Pages of the Mobile Application

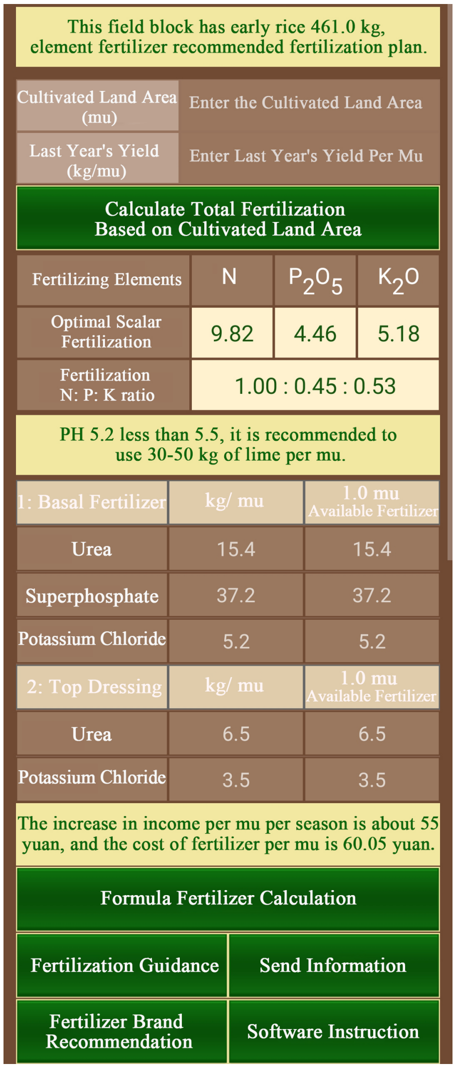
(a) Acquire Soil Analysis by GPS Tracking


(b) Acquire Soil Analysis by Selecting Places

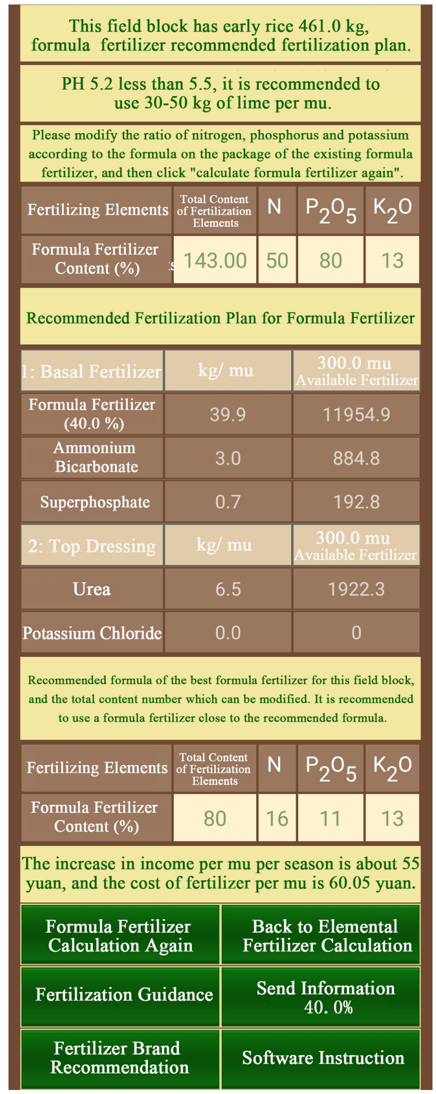

*Note:* The left panel shows that farmers can acquire soil testing data by GPS tracking or by choosing locations. The app then displays the amount of pH value, organic matter, nitrogen, phosphorus, and potassium in farmers' plots. The right panel shows a set of locations that farmers can select from.

Figure A.4: Dynamic Fertilizer Recommendations

(a) Combination of Different Individual Fertilizers



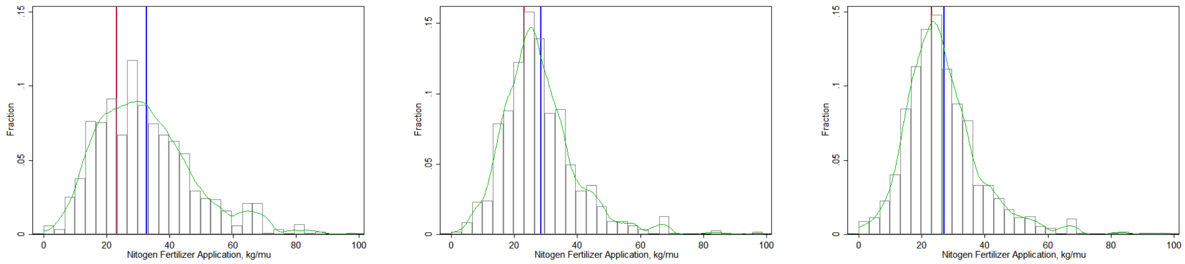

(b) Combination of Compound + Other Individual Fertilizers




*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 A.5: Total Nitrogen Application [Used >> Recommended]

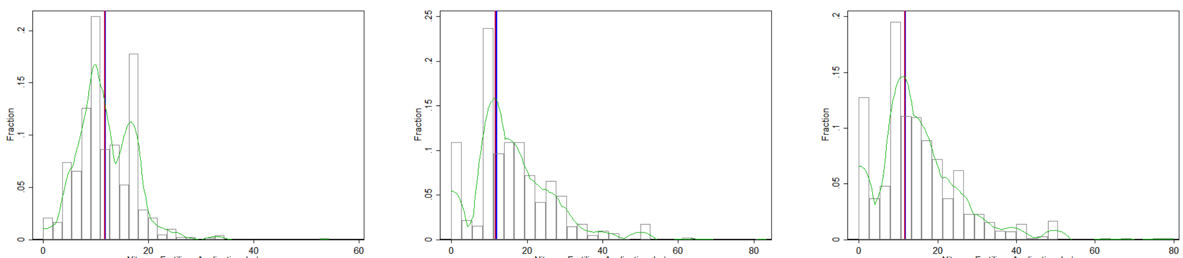
(a) Before Interventions, 2019 (b) After 1st Intervention, 2020 (c) After 2nd Intervention, 2021




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 A.6: Nitrogen Application in the Planting Stage [Used ≈ Recommended]

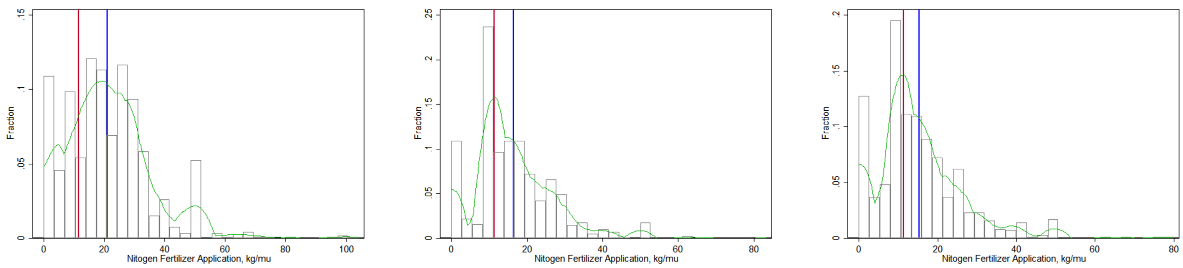
(a) Before Interventions, 2019 (b) After 1st Intervention, 2020 (c) After 2nd Intervention, 2021




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 A.7: Top-dressing Nitrogen Application in the Growing Stage [Used >> Recommended]

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

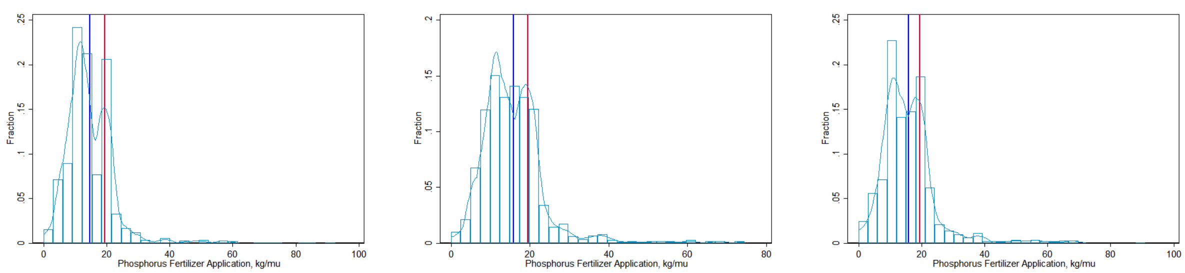



Note: This figure shows the distribution of nitrogen fertilizer application (kg/mu) in the growing stages. The red line indicates the mean of nitrogen recommendations, while the blue line shows the mean of actual use. The figure shows a clear pattern that nitrogen fertilizers are over-applied during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

Figure A.8: Total Phosphorus Application [Used < Recommended]

(Red = Mean Recommendations, blue = Mean Application)

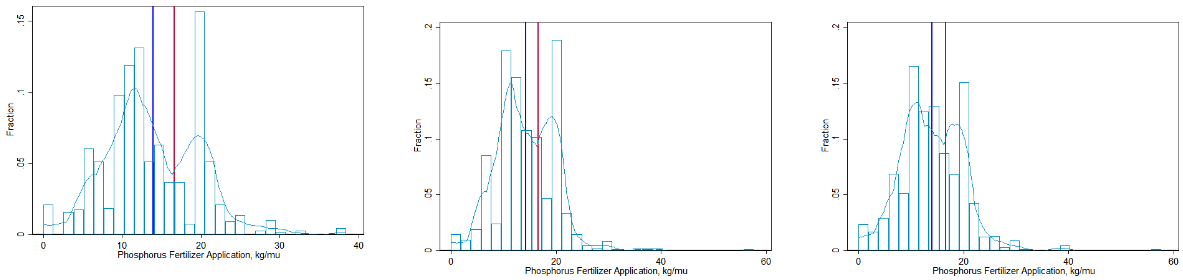
Before Interventions 2019 After Phase 1 Intervention, 2020 After Phase 2 Intervention, 2021




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 A.9: Phosphorus Application in the Planting Stage [Used < Recommended]

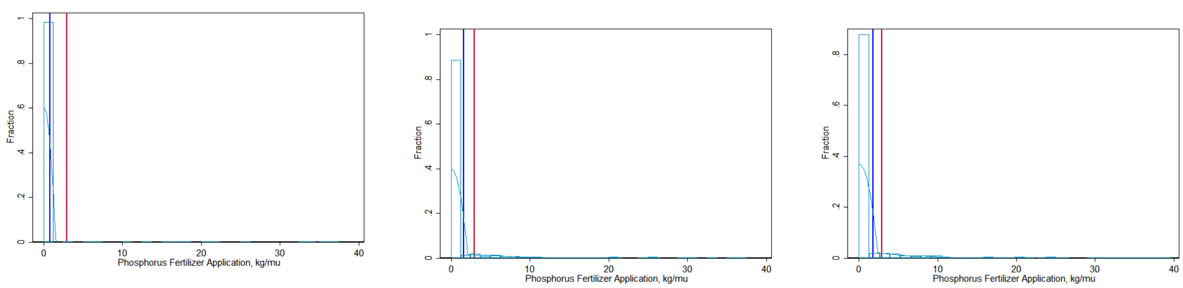
(a) Before Interventions, 2019 | (b) After Phase 1 Intervention, 2020 | (c) After Phase 2 Intervention, 2021




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 A.10: 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

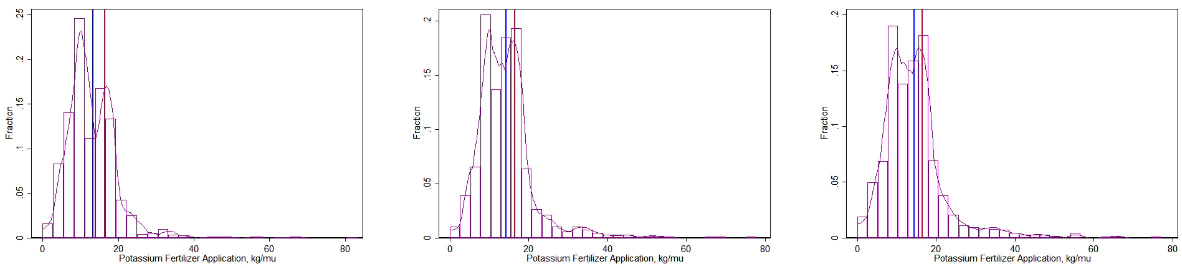



Note: This figure shows the distribution of phosphorus fertilizer application (kg/mu) in the growing stages. The red line indicates the mean of phosphorus recommendations, while the blue line shows the mean of actual use. The figure suggests underuse of phosphorus in the growing stages, and our interventions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

Figure A.11: Total Potassium Application [Used < Recommended]

(Red = Mean Recommendations, blue = Mean Application)

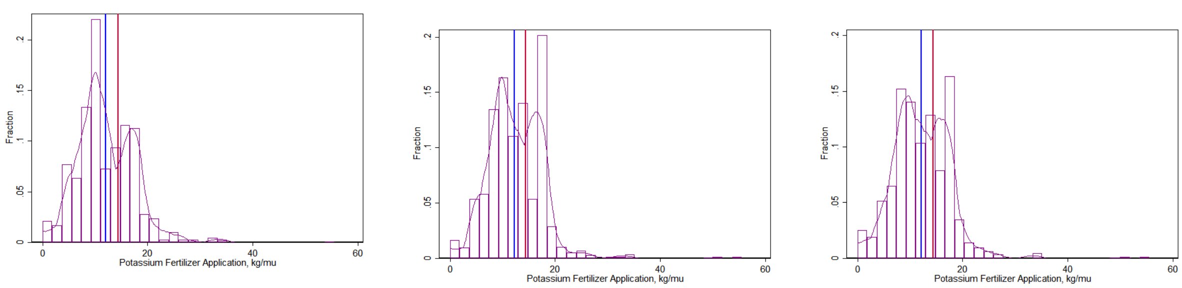
Before Interventions 2019 | After Phase 1 Intervention, 2020 | After Phase 2 Intervention, 2021




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 A.12: Potassium Application in the Planting Stage [Used < Recommended]

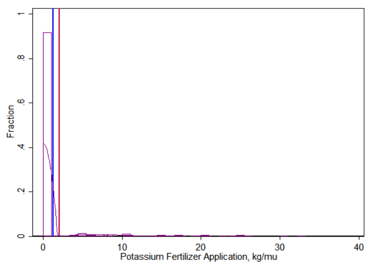
(a) Before Interventions 2019 | (b) After Phase 1 Intervention, 2020 | (c) After Phase 2 Intervention, 2021



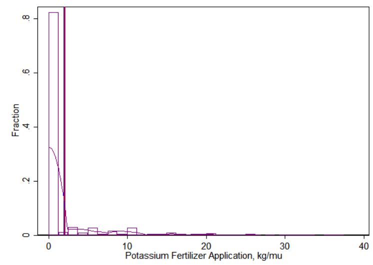

Note: This figure shows the distribution of potassium fertilizer application (kg/mu) in the planting stages. The red line indicates the mean of potassium recommendations, while the blue line shows the mean of actual use. The figure suggests underuse of potassium in the planting stages.

Figure A.13: Top-dressing Potassium Application in the Growing Stage [Used < Recommended]

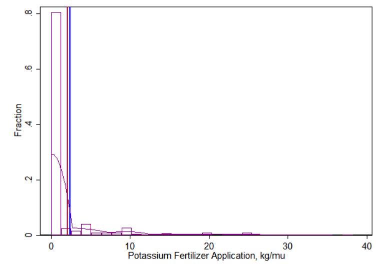
(a) Before Interventions 2019




(b) After Phase 1 Intervention, 2020



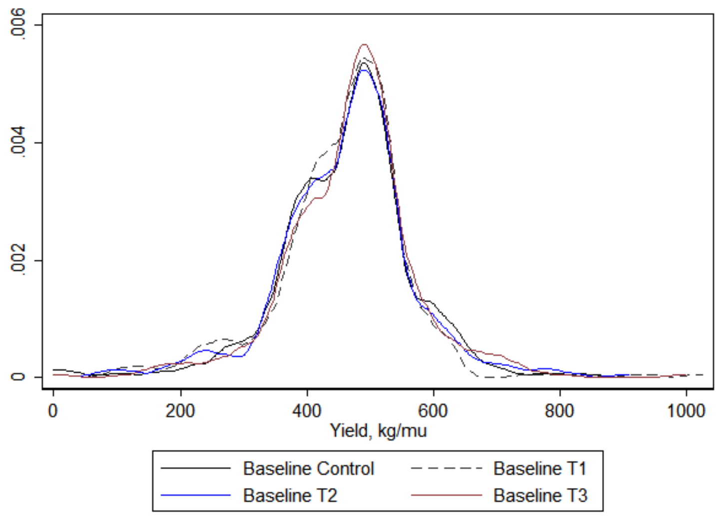

(c) After Phase 2 Intervention, 2021




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 A.14: Yields before/after the First-phase Interventions

(a) Yields at Baseline (kg/mu)




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

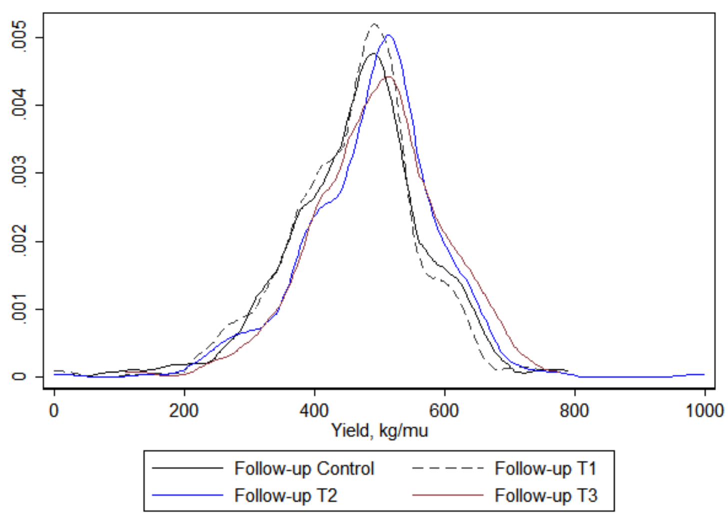


Table A.1: Distance and Fertilizer Gap between Applications and Recommendations

| | (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 A.2: 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 (1.6) and (1.7). 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 A.3: IV Estimation II: Using T2 and T3 Indicators as IVs for Each Fertilizer Gap

| Dept. Vars. | (1) Yields | (2) Yields | (3) Yields | (4) Yields | (5) Yields | (6) 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 A.4: IV Estimation: Deviation in Fertilizer Application and Yields

| | (1)<br>IV-First Stage<br>$logGap^2$ | (2)<br>2SLS<br>Yields | (3)<br>2SLS<br>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 1.9, 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 (1.6) and (1.7). 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 A.5: IV Estimation: Deviation in Fertilizer Application and Revenues/Costs

| | (1) IV-First Stage $logGap^2$ | (2) 2SLS Revenues | (3) 2SLS Fertilizer Cost | (4) 2SLS 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 1.10, 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 (1.6) and (1.7). 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%.

## A.2 Proofs.

In this section we detail the derivation described in Section 1.4. We begin with equation (1.4).

After observing the profit $\pi_t$ generated by the realized states $b_{Nt}$ and $b_{Kt}$, the farmer believes that the effectiveness of nitrogen fertilizer is $b_{Nt}$ and the effectiveness of phosphorus (P)/potassium (K) fertilizers is $\tilde{b}_{Kt}$, which satisfies:

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

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

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

The interpretation for this equation is that the farmer believes that whatever action she chooses, she can infer a unique signal $\tilde{b}_K$ about $\Theta_K$ to equalize her belief and the real profit. At $\Pi = \tilde{\Pi}$, we could derive the distorted $\tilde{b}_K$ as:

$$\tilde{b}_K\left(a_N, a_K, b_K\right) = \tilde{b}_K\left(a_K, b_K\right) = \log\left(C + f_2\left(a_K\right)\exp\left(b_K\right)\right) - \log\left(f_2\left(a_K\right)\right) \tag{A.1}$$

where $C = (1-\lambda) f_1\left(a_N\right)\exp\left(b_N\right)$.

Turning back to the true profit function in Equation (1.1), we define $a_K^*$ as the optimal phosphorus (P)/potassium (K) fertilizer input. By taking partial derivative of the true profit function in Equation (1.1) w.r.t. $a_K$, $a_K^*$ then satisfies: $f_2'\left(a_K^*\right)\exp\left(\theta_K + \sigma_K^2/2\right) = c_K$. Take logs on both sides and rearrange terms, we could obtain $\log\left(f_2'\left(a_K^*\right)\right) = \log\left(c_K\right) - \theta_K - \sigma_K^2/2$, that is,

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

Similarly, for the the misspecified model of the profit function in Equation (1.3), we take the derivative of the misspecified profit function w.r.t. $\tilde{a}_K$ and obtain the following relationship:

$$\log\left(f_2'\left(\tilde{a}_K^*\right)\right) = \log\left(c_K\right) - \tilde{\theta}_K - \sigma_K^2/2 \tag{A.3}$$

where $\tilde{a}_K^*$ is the the optimal phosphorus (P)/potassium (K) fertilizer input under the misspeciied profit function. In Equation (A.1) that $\tilde{b}_K\left(a_K, b_K\right) = \log\left(C + f_2\left(a_K\right)\exp\left(b_K\right)\right) - \log\left(f_2\left(a_K\right)\right)$, we have $C < 0$ because $\lambda > 1$ and $f_1\left(a_N\right) > 0$. Thus then we can always find a negative number $A \in \mathbb{R}^-$ to ensure the following replacement,

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

Then we get

$$\begin{aligned}\tilde{b}_K\left(a_K, b_K\right) &= \log\left(C + f_2\left(a_K\right)\exp\left(b_K\right)\right) - \log f_2\left(a_K\right) \\ &= A + \log\left(f_2\left(a_K\right)\exp\left(b_K\right)\right) - \log f_2\left(a_K\right) = A + b_K\end{aligned} \tag{A.4}$$

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

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

where $b_{t+1}$ is the effectiveness of fertilizer at time $t + 1$. The above expectation is the one of an outside observer who knows the true fundamental $\Theta$ and the agent's subjective belief $\tilde{b}$, so that, being able to deduce the fertilizer input $a^*(t + 1, \tilde{\theta}_t)$, she knows the distribution of $\tilde{b}_{t+1}$.The function can be thought of as the agent's mean surprise regarding the fundamental. We then redefine how are farmer's subjective beliefs updated in the limit as:

$$g_K(x) = \lim_{t\to\infty} g_K(t, x) = \mathbb{E}\left[\tilde{b}_K\left(b_{K,t+1}, a_K^*(x)\right)\right] - x$$

In our context, farmers need to learn about phosphorus (P)/potassium (K) fertilizer. According to this expectation of the updating convergence function, we obtain:

$$g_K(\tilde{\theta}_K) = \lim_{t\to\infty} E\left(\tilde{b}_K\left(b_K, a_K^*(\tilde{\theta}_K)\right)\right) - \tilde{\theta}_K \tag{A.5}$$

Since $\tilde{b}_K\left(a_K, b_K\right) = A + b_K$ in Equation (A.4), as $t \to \infty$, $g_K(\tilde{\theta}_K) = A + E\left(b_K\right) - \tilde{\theta}_K$, and from Equation (A.3) 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 $g_K$ function in Equation (A.5), and solve $g_K(\tilde{\theta}_K) = 0$, we acquire:

$$A + \theta_K - \tilde{\theta}_K = 0$$

where $\theta_K$ is defined as the mean of $b_K$ in Section 1.4. Therefore, we have:

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

Based on Equation (A.2) that $\theta_K = \log\left(c_K\right) - \log\left(f_2'\left(a_K^*\right)\right) - \sigma_K^2/2$, we obtain:

$$\log\left(f_2'\left(\tilde{a}_K^*\right)\right) = \log\left(c_K\right) - \theta_K - A - \sigma_K^2/2 = \log\left(f_2'\left(a_K^*\right)\right) - A \tag{A.6}$$

Since $A < 0$, so we have $\log\left(f_2'\left(\tilde{a}_K^*\right)\right) > \log\left(f_2'\left(a_K^*\right)\right)$, and hence then $f_2'\left(\tilde{a}_2^*\right) \geq f_2'\left(a_K^*\right)$.

Due to the concavity of the functions $f_1$ and $f_2$, we could derive that,

$$\tilde{a}_K^* < a_K^* \tag{A.7}$$

Recall that $\tilde{a}_K^*$ is the optimal phosphorus (P)/potassium (K) fertilizer input under the misspecified model and $a_K^*$ represents the optimal phosphorus (P)/potassium (K) fertilizer input under the true model.

As for nitrogen fertilizer use, by taking the first order condition of Equation (1.1) and (1.3) with respect to $a_N$, 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 defined as the optimal nitrogen use under the misspecifed model, and $a_N^*$ is defined as the optimal nitrogen application under the true model. Obviously, we can get:

$$a_N'^* > a_N^* \tag{A.8}$$

**Prediction 4.** *$a_N'^* > a_N^*$, $\tilde{a}_K^* < a_K^*$, 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/potassium fertilizers at equilibrium under the misspcified and true models, respectively.*

**Prediction 5.** *If farmers' fertilizer input $a_N$ decreases and moves closer to $a_N^*$ and $\tilde{a}_K$ increases and moves toward $a_K^*$ in the current period, then their belief about the effectiveness of phosphorus (P)/potassium (K) fertilizer $\tilde{b}_K$ will also increase and rise up toward $b_k$. To be specific, if farmers are nudged to adopt less nitrogen fertilizer and more phosphorus (P)/potassium (K) fertilizers in their actions in the current period, then their subjective beliefs $\tilde{b}_K$ about the effectiveness of P/K in the next period will move upwards toward the true value o the effectiveness. As such, farmers' undervaluation of P/K will decrease.*

*Proof.*

$$\begin{aligned}\tilde{b}_K\left(a_N', a_K', b_K\right) &= \log\left(C' + f_2\left(a_K'\right)\exp\left(b_K\right)\right) - \log\left(f_2\left(a_K'\right)\right)\\ &= log\left(\frac{(1-\lambda)f_1(a_N')exp(b_N)}{f_2(a_N')} + exp(b_K)\right)\end{aligned} \tag{A.9}$$

Define $a_N'$ and $a_K'$ are the input level at some certain period, and $a_N''$ and $a_k''$ are different input levels. Obviously, $\tilde{b}_K\left(a_N', a_K', b_K\right)$ is increasing as $a_N'$ decreases and $a_K'$ increases. If $a_N'' > a_N' > a_N^*$, $a_K'' < a_K' < a_K^*$, then we'll have $\tilde{b}_K\left(a_N', a_K', b_K\right) > \tilde{b}_K\left(a_N'', a_K'', b_K\right)$. □

**Prediction 6.** *Correcting the overestimation/misspecification of the return to nitrogen fertilizer leads to a lower level of nitrogen fertilizer input $a_N$ directly and a higher application of phosphorus (P)/potassium (K) fertilizer $a_K$. Farmers' belief about the effectiveness of phosphorus (P)/potassium*

*(K) fertilizers will gradually move toward the true realized state. 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.*

**Proof of Prediction (3)** Based on equation (A.1), if $\lambda$=1, then $C = (1-\lambda) f_1(a_N) \exp(b_N) = 0$

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

**Proof of Convergence**
We prove convergence by verifying the conditions specified in Assumption 1 in Heidhues, Kőszegi, and Strack (2021).

(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 $a_K$.

$$|b_K - \tilde{b}_K(b_k, a_N, a_K)| =$$
$$|log(\frac{C}{exp(b_K)f_2(a_K)} + 1)|$$

We need $\frac{(1-\lambda)f_1(a_N)\exp(b_N)}{exp(b_K)f_2(a_K)}$ to be bounded away from -1. In other words, there exist $\epsilon > 0$ such that for all feasible $a_N$, $b_N$, $a_K$, $b_K$,

$$\frac{(1-\lambda) f_1(a_N)\exp(b_N)}{exp(b_K)f_2(a_K)} > -1 + \epsilon \tag{A.10}$$

Equation (A.10) indicates that the output component of phosphorus (P)/potassium (K) should not be too smaller than output component of nitrogen (N). For instance, if $\lambda = 2$, then phosphorus (P)/potassium (K) and nitrogen (N) have equivalent contributions to the total output.

(II) $\frac{\partial \tilde{b}_K(a_N, a_K, b_K)}{\partial a_K}$ is bounded.

Since we can rewrite the expression of $\tilde{b}_K(a_N, a_K, b_K)$ as follows

$$\tilde{b}_K(a_N, a_K, b_K) = log(\frac{C}{f_2(a_K)} + exp(b_K))$$

If $\frac{(1-\lambda)f_1(a_N)\exp(b_N)}{exp(b_K)f_2(a_K)}$ is bounded away from -1, then $f_2(a_K)$ must be positive and large enough. As long as $f_2'(a_K) < \infty$, equation A.10 suffices to ensure that the derivative with regard to $a_K$ is bounded.

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

Recall that according to the definition of $a^*_K(t, \tilde{\theta}_K)$ and $a^*_K(\tilde{\theta}_K)$ can be rewritten as follows:

$$a^*_K(t, \tilde{\theta}_K) = argmax_a \int [\int \tilde{\Pi}(a_{Kt}, \tilde{b}_{Kt})\phi(\tilde{b}_{Kt}; x, \sigma^2)d\tilde{b}_{Kt}]\phi(x; \tilde{\theta}_K, v_{t-1})dx$$

$$a^*_K(\tilde{\theta}_K) = argmax_a \int \tilde{\Pi}(a_{Kt}, \tilde{b}_{Kt})\phi(\tilde{b}_{Kt}; \tilde{\theta}_K, \sigma^2)d\tilde{b}_{Kt}$$

where $\tilde{\Pi}(a_{Kt}, \tilde{b}_{Kt})$ is the abbreviation of $\tilde{\Pi}(a_{Nt}, b_{Nt}, a_{Kt}, \tilde{b}_{Kt})$. $\phi(\tilde{b}_{Kt}; x, \sigma^2)$ refers to the density function of $\tilde{b}_{Kt}$, with the mean being $x$ and variance being $\sigma^2$.

As $\frac{\partial^2 \tilde{\Pi}(a_{Kt}, \tilde{b}_{Kt}}{\partial a_{Kt} \partial \tilde{b}_{Kt}}$ is continuous in our setting, $a^*_K(t, \tilde{\theta}_K)$ and $a^*_K(\tilde{\theta}_K))$ can be implicitly defined as follows:

$$f_2'(a^*_K(\tilde{\theta}_K))) \int exp(\tilde{b}_{Kt})\phi(\tilde{b}_{Kt}; \tilde{\theta}_K, \sigma^2)d\tilde{b}_{Kt} = c_K \tag{A.11}$$

and

$$f_2'(a^*_K(t, \tilde{\theta}_K)) \int [\int exp(\tilde{b}_{Kt})\phi(\tilde{b}_{Kt}; x, \sigma^2)d\tilde{b}_{Kt}]\phi(x; \tilde{\theta}_K, v_{t-1})dx = c_K \tag{A.12}$$

Equation A.11 and A.12 illustrate that $\int exp(\tilde{b}_{Kt})\phi(\tilde{b}_{Kt}; \tilde{\theta}_K, \sigma^2)d\tilde{b}_{Kt}$ can be approximated by $\int [\int exp(\tilde{b}_{Kt})\phi(\tilde{b}_{Kt}; x, \sigma^2)d\tilde{b}_{Kt}]\phi(x; \tilde{\theta}_K, v_{t-1})d$, then $f_2'(a^*_K(\tilde{\theta}_K))$ can be approximated by $f_2'(a^*_K(t, \tilde{\theta}_K))$. If the inverse of $f_2'(.)$ is well behaved, then $a^*_K(\tilde{\theta}_K)$ can be approximated by $a^*_K(t, \tilde{\theta}_K)$. More precisely, if there exists $\epsilon > 0$ such that for any $a_{Kt}$,

$$|f_2''(a_{Kt})| \geq \epsilon \tag{A.13}$$

then it would be possible to complete the aforementioned reasoning of approximation. Below we prove that the condition A.13 is sufficient.

Observe that expression $\int exp(\tilde{b}_{Kt})\phi(\tilde{b}_{Kt}; x, \sigma^2)d\tilde{b}_{Kt}$ can be simplified as follows:

$$\begin{aligned} intexp(\tilde{b}_{Kt})\phi(\tilde{b}_{Kt}; x, \sigma^2)d\tilde{b}_{Kt} = \\ \int \frac{1}{\sigma\sqrt{2\pi}} exp(\tilde{b}_{Kt}) exp(-\frac{(\tilde{b}_{Kt} - \tilde{\theta}_K)^2}{2\sigma^2})d\tilde{b}_{Kt} = \\ exp(\tilde{\theta}_K + \frac{\sigma^2}{2}) \end{aligned}$$

We have

$$\left|a_K^*(t,\tilde{\theta}_K) - a_K^*(\tilde{\theta}_K)\right| =$$
$$exp(\frac{\sigma^2}{2})\int[exp(\tilde{\theta}_K)) - exp(x)]\frac{1}{\sqrt{2\pi}v_{t-1}}exp(-\frac{(x-\tilde{\theta}_K)^2}{2v_{t-1}^2})dx$$

Let $\varphi = \frac{x-\tilde{\theta}_K}{v_{t-1}}$, the expression above can be rewritten as follows:

$$\left|a_K^*(t,\tilde{\theta}_K) - a_K^*(\tilde{\theta}_K)\right| =$$
$$\frac{\sigma^2}{2\sqrt{2\pi}}exp(\tilde{\theta}_K)\int[1 - exp(v_{t-1}\varphi)]exp(-\frac{\varphi^2}{2})d\varphi$$

We claim that $\frac{1}{\sqrt{2\pi}}\int[1-exp(v_{t-1}\varphi)]exp(-\frac{\varphi^2}{2})d\varphi$ is of order $O(\frac{1}{t^2})$. To see this, consider simplifying this expression:

$$\frac{1}{\sqrt{2\pi}}\int[1 - exp(v_{t-1}\varphi)]exp(-\frac{\varphi^2}{2})d\varphi =$$
$$1 - \frac{1}{\sqrt{2\pi}}\int exp(v_{t-1}\varphi)exp(-\frac{\varphi^2}{2})d\varphi =$$
$$1 - \frac{1}{\sqrt{2\pi}}\int exp(-\frac{(\varphi - v_{t-1})^2}{2} + \frac{v_{t-1}^2}{2})d\varphi =$$
$$1 - exp(\frac{v_{t-1}^2}{2})$$

Since $v_{t-1} = \frac{1}{v_0^{-1}+(t-1)\sigma^{-2}}$ is of order $O(\frac{1}{t})$, we have

$$\frac{1}{\sqrt{2\pi}}\int[1 - exp(v_{t-1}\varphi)]exp(-\frac{\varphi^2}{2})d\varphi =$$
$$1 - exp(\frac{v_{t-1}^2}{2}) =$$
$$1 - 1 + O(\frac{v_{t-1}^2}{2}) =$$
$$O(\frac{v_{t-1}^2}{2}) = O(\frac{1}{t^2})$$

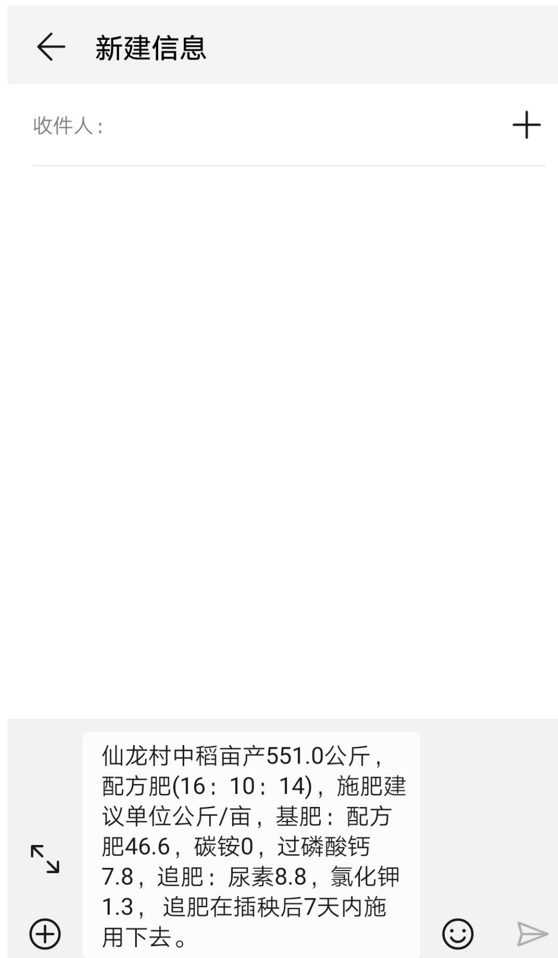
Thus Assumption (III) is satisfied when condition A.13 holds.

## A.3 Documentations about the Implement of Experiments

Figure A.15: 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 A.15a shows that farmers can choose a specific fertilizer producer to order customized fertilizers by inputting names, contacts and coordinates of the plots. Figure A.15b shows that the app can generate a message with fertilizer recommendations to non-smart phone users.

# Appendix B

# Chapter 2 Appendix

## B.1 Supplementary Tables and Figures

Table B.1: Distribution of Students' Preferences and Admitted Majors

| | Economics / Business | Science | Law/ Sociology | Humanities/ Languages | Total |
|---|---|---|---|---|---|
| Panel A: Overall Preferences in Different Ranks | | | | | |
| 1st Preference | 989 | | | | 989 |
| 2nd Preference | 798 | 78 | 73 | 34 | 989 |
| 3rd Preference | 687 | 71 | 112 | 81 | 989 |
| 4th Preference | 570 | 98 | 163 | 94 | 989 |
| 5th Preference | 505 | 110 | 148 | 110 | 989 |
| 6th Preference | 442 | 107 | 143 | 93 | 989 |
| Admitted to the Major | 493 | 197 | 176 | 123 | 989 |
| Panel B: Students Admitted to Their Ranked-preference Majors | | | | | |
| Rank | 1 st | 2nd | 3rd | 4th | |
| Count | 243 | 229 | 173 | 108 | 989 |

This table describes the distribution of preferred majors and ultimate admissions among 989 students by matching the survey data to the university administrative admission database. Panel A quantifies the distribution of applicants whose 1st, 2nd,..., 6th preferences fell into the categories for economics/business majors, natural sciences, law/sociology, humanities/language studies. Panel B illustrates which major (e.g., first choice, second choice) individuals were ultimately admitted into.

Table B.2: Proportion of Students Taking Compulsory Courses: Concepts in Economics

| Course | Micro | | Macro | | Finance | |
|---|---|---|---|---|---|---|
| Semester | Econ | Non-econ | Econ | Nonecon | Econ | Eonecon |
| 1 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| 2 | 100.00% | 40.70% | 0.00% | 0.00% | 0.00% | 0.00% |
| 3 | 100.00% | 40.70% | 100.00% | 37.63% | 61.26% | 15.34% |
| 4 | 100.00% | 40.70% | 100.00% | 37.63% | 65.98% | 15.34% |
| 5 | 100.00% | 40.70% | 100.00% | 37.63% | 65.98% | 15.34% |
| 6 | 100.00% | 40.70% | 100.00% | 37.63% | 65.98% | 15.34% |
| 7 | 100.00% | 40.70% | 100.00% | 37.63% | 65.98% | 15.34% |
| 8 | 100.00% | 40.70% | 100.00% | 37.63% | 65.98% | 15.34% |

Table B.3: Proportion of Students Taking Compulsory Courses: Concepts in Statistics

| Course | Probability | | Statistics | | Econometrics | |
|---|---|---|---|---|---|---|
| Semester | Econ | Non-econ | Econ | Non-econ | Econ | Nonecon |
| 1 | 0.00% | 0.00% | 4.57% | 11.45% | 0.00% | 0.00% |
| 2 | 15.28% | 11.45% | 4.57% | 11.45% | 0.00% | 0.00% |
| 3 | 100.00% | 53.99% | 4.57% | 11.45% | 0.00% | 0.00% |
| 4 | 100.00% | 53.99% | 56.69% | 23.72% | 32.28% | 0.00% |
| 5 | 100.00% | 53.99% | 70.08% | 23.72% | 71.34% | 20.86% |
| 6 | 100.00% | 53.99% | 70.08% | 23.72% | 71.34% | 20.86% |
| 7 | 100.00% | 53.99% | 70.08% | 23.72% | 71.34% | 20.86% |
| 8 | 100.00% | 53.99% | 70.08% | 23.72% | 71.34% | 20.86% |

Table B.4: Pre-college Rankings and Decision-making

| Dep. Var. | (1) MPL1 Risk Neutral | (2) MPL2 Risk Neutral | (3) Law of Large Numbers | (4) Two Indifferent Choices | (5) Exact Representiveness | (6) Dictator's Sharing | (7) Bystander's Belief |
|---|---|---|---|---|---|---|---|
| Non-Economics× Top Ranking | -0.133*** | -0.087** | -1.213** | 0.521 | -1.031* | -7.000 | 27.165* |
| | (0.040) | (0.040) | (0.532) | (0.436) | (0.546) | (15.917) | (16.001) |
| Non-Economics× Middle Ranking | -0.118** | -0.030 | -1.412** | -0.350 | -1.343** | 11.401 | 18.771 |
| | (0.047) | (0.048) | (0.637) | (0.521) | (0.653) | (18.870) | (19.697) |
| Non-Economics× Bottom Ranking | -0.103** | 0.006 | -1.009 | 0.689 | -1.124 | 0.859 | 17.671 |
| | (0.050) | (0.051) | (0.677) | (0.554) | (0.694) | (20.366) | (21.678) |
| Common Support of Major Preference | X | X | X | X | X | X | X |
| Constant | 0.353*** | 0.312*** | 18.314*** | 15.092*** | 25.443*** | 187.892*** | 165.262*** |
| | (0.026) | (0.027) | (0.354) | (0.290) | (0.363) | (10.279) | (10.624) |
| Observations | 802 | 802 | 802 | 802 | 802 | 274 | 275 |
| R-squared | 0.043 | 0.019 | 0.022 | 0.013 | 0.023 | 0.053 | 0.019 |

In this table, we introduce three interaction terms between pre-college rankings and a dummy of non-economics education. The first row reports the differences in the treatment effects between top-ranking (pre-college) students who are assigned to a non-economics majors, compared to the average outcomes of economics students, while the second and third show the differences in treatments of middle-ranking and bottom-ranking students, relative to the average outcomes of economics students.

The dependent variables in columns (1) and (2) are the share of risk neutral students in MPL 1 and MPL 2, in which we pool WTA and WTP together. Columns (3), (4), and (5) report results on probabilistic belief questions on the law of large numbers (LLN), two identical choices, and Exact Representativeness (ER), respectively. The outcome variables in columns (6) and (7) are the Dictators' actual sharing and Bystander's beliefs regarding the Dictators' sharing in the Dictator Game.

All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics major category (*Major-Preference FX*), and limit the regression sample to students who put both economics and non-economics majors in their rank-order list (*Common Support of Major Preference*).

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table B.5: The Rank of the Admitted Major and Decision-making

| Dep. Var. | (1) MPL1 Risk Neutral | (2) MPL2 Risk Neutral | (3) Law of Large Numbers | (4) Two Indifferent Choices | (5) Exact Representiveness | (6) Dictator's Sharing | (7) Bystander's Belief |
|---|---|---|---|---|---|---|---|
| Admitted in 3rd | 0.031 | -0.077 | 0.407 | -0.786 | 0.214 | 10.476 | 18.890 |
| | (0.055) | (0.050) | (0.732) | (0.601) | (0.750) | (20.406) | (23.570) |
| Admitted in 4th | 0.010 | -0.036 | 0.659 | -1.188* | 0.589 | -24.469 | 49.618* |
| | (0.063) | (0.056) | (0.839) | (0.689) | (0.860) | (23.028) | (27.543) |
| Admitted in 5th | 0.040 | 0.009 | 0.541 | -1.194 | -1.828* | 7.373 | 13.042 |
| | (0.072) | (0.064) | (0.964) | (0.790) | (0.987) | (32.910) | (28.412) |
| Admitted in 6th | -0.017 | -0.044 | 2.410** | -0.935 | 0.102 | 1.485 | 10.245 |
| | (0.083) | (0.077) | (1.115) | (0.914) | (1.142) | (29.531) | (42.952) |
| Admitted in 6th+ | X | X | X | X | X | X | X |
| F-Joint Test | 0.17 | 1.21 | 0.95 | 0.99 | 1.32 | 0.49 | 0.78 |
| *P-value* | (0.97) | (0.30) | (0.45) | (0.42) | (0.25) | (0.78) | (0.56) |
| Econ-related Majors FX | X | X | X | X | X | X | X |
| Major-Preference FX | X | X | X | X | X | X | X |
| Common Support of Major Preference | X | X | X | X | X | X | X |
| Constant | 0.269*** | 0.314*** | 17.150*** | 15.881*** | 24.844*** | 188.856*** | 164.705*** |
| | (0.030) | (0.025) | (0.401) | (0.329) | (0.411) | (11.933) | (12.100) |
| Observations | 802 | 802 | 802 | 802 | 802 | 274 | 275 |
| R-squared | 0.044 | 0.019 | 0.028 | 0.015 | 0.031 | 0.059 | 0.033 |

This table reports the results using regression equation (4). Variables *Admitted in 3rd - Admitted in 6th* denote that students were admitted to the third,...,sixth rank of their rank-order list. *Admitted in 6th+* indicates that students didn't meet the cutoffs of all the six majo-preferences and were assigned to a major by CUFE.

The dependent variables in columns (1) and (2) are the share of risk neutral students in MPL 1 and MPL 2, in which we pool WTA and WTP together. Columns (3), (4), and (5) report results on probabilistic belief questions on the law of large numbers (LLN), two identical choices, and Exact Representativeness (ER), respectively. The outcome variables in columns (6) and (7) are the Dictators' actual sharing and Bystander's beliefs regarding the Dictators' sharing in the Dictator Game.

All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics major category (*Major-Preference FX*), and limit the regression sample to students who put both economics and non-economics majors in their rank-order list (*Common Support of Major Preference*).

F-test for the joint significance of the dummies for the position of admitted major in the rank-order list and p-values are provided.

Table B.6: Out of Pocket, Preferences and Beliefs

| Dep. Var. | (1) MPL1 Risk Neutral | (2) MPL2 Risk Neutral | (3) Law of Large Numbers | (4) Two Indifferent Choices | (5) Exact Representiveness | (6) Dictator's Sharing | (7) Bystander's Belief |
|---|---|---|---|---|---|---|---|
| Econ | 0.154*** | 0.076** | 1.519*** | -0.496 | 1.411*** | -10.276 | -17.631 |
| | (0.032) | (0.033) | (0.430) | (0.353) | (0.444) | (12.575) | (13.075) |
| In Deficit | 0.002 | 0.002 | -0.013 | 0.002 | -0.003 | 0.180 | 6.583 |
| | (0.001) | (0.001) | (0.015) | (0.013) | (0.016) | (0.268) | (5.178) |
| Common Support of Major Preference | X | X | X | X | X | X | X |
| Constant | 0.218*** | 0.250*** | 16.946*** | 15.460*** | 24.173*** | 192.755*** | 186.403*** |
| | (0.020) | (0.020) | (0.266) | (0.218) | (0.274) | (8.047) | (8.199) |
| Observations | 796 | 796 | 796 | 796 | 796 | 270 | 275 |
| R-squared | 0.031 | 0.009 | 0.016 | 0.002 | 0.013 | 0.004 | 0.012 |

In this table, we introduce a variable of financial status last month, *In Deficit*, as a control variable, which equals the difference between money from family and spending last month.

The dependent variables in columns (1) and (2) are the share of risk neutral students in MPL 1 and MPL 2, in which we pool WTA and WTP together. Columns (3), (4), and (5) report results on probabilistic belief questions on the law of large numbers (LLN), two identical choices, and Exact Representativeness (ER), respectively. The outcome variables in columns (6) and (7) are the Dictators' actual sharing and Bystander's beliefs regarding the Dictators' sharing in the Dictator Game.

All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics major category (*Major-Preference FX*), and limit the regression sample to students who put both economics and non-economics majors in their rank-order list (*Common Support of Major Preference*).

Standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B.7: Heterogeneous Effects by Gender

| Dep. Var. | (1) MPL1 Risk Neutral | (2) MPL2 Risk Neutral | (3) Law of Large Numbers | (4) Identical Choices | (5) Exact Representiveness | (6) Dictator's Sharing | (7) Bystander's Belief |
|---|---|---|---|---|---|---|---|
| Econ* Male | 0.111* | 0.015 | 1.850** | -0.620 | -1.237 | -15.266 | -36.177 |
| | (0.066) | (0.068) | (0.885) | (0.735) | (0.914) | (26.811) | (27.031) |
| Econ | 0.084** | 0.047 | 0.579 | -0.072 | 1.667*** | 4.217 | -7.702 |
| | (0.043) | (0.044) | (0.571) | (0.474) | (0.590) | (16.437) | (18.144) |
| Male | 0.084** | 0.063 | 1.130** | 0.693 | 1.804*** | -10.453 | 8.045 |
| | (0.042) | (0.043) | (0.563) | (0.468) | (0.581) | (17.834) | (17.292) |
| Major-Preference FX | X | X | X | X | X | X | X |
| Common Support of Major Preference | X | X | X | X | X | X | X |
| Constant | 0.199*** | 0.237*** | 16.653*** | 15.153*** | 23.628*** | 191.426*** | 184.770*** |
| | (0.026) | (0.026) | (0.342) | (0.284) | (0.354) | (10.040) | (11.302) |
| Observations | 802 | 802 | 802 | 802 | 802 | 274 | 275 |
| R-squared | 0.065 | 0.020 | 0.049 | 0.011 | 0.035 | 0.056 | 0.026 |

This table presents the average difference in survey responses between economics students and non-economics students by gender by introducing an interaction term *Male* × *Economics* using regression equation (3).
The dependent variables in columns (1) and (2) are the share of risk neutral students in MPL 1 and MPL 2, in which we pool WTA and WTP together. Columns (3), (4), and (5) report results on probabilistic belief questions on the law of large numbers (LLN), two identical choices, and Exact Representativeness (ER), respectively. The outcome variables in columns (6) and (7) are the Dictators' actual sharing and Bystander's beliefs regarding the Dictators' sharing in the Dictator Game.
All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics major category (*Major-Preference FX*), and limit the regression sample to students who put both economics and non-economics majors in their rank-order list (*Common Support of Major Preference*).
Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table B.8: Robust Check Using [-0.15, 0.15] Times the Standard Deviation

| | Risk Preferences | | Probabilistic Beliefs (3)-(5) | | | Dictator Game | | Trust Game (8)-(11) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| | MPL1 | MPL2 | Law of Large | Two Indifferent | Exact | Dictator's | Bystander's | Proposer's | Bystander's | Reciprocity of | Bystander's Belief |
| Dep. Var. | Risk Neutral | Risk Neutral | Numbers | Choices | Representiveness | Sharing | Belief | Sharing | Belief | Player 2 | about Reciprocity |
| Econ=1 | 0.147*** | 0.070** | 1.439*** | -0.402 | 1.409*** | -8.551 | -15.888 | -11.689** | 9.788 | -3.370 | 0.721 |
| | (0.031) | (0.032) | (0.410) | (0.339) | (0.422) | (12.038) | (12.763) | (5.768) | (6.917) | (3.530) | (5.965) |
| M′=(-50, 0, 50) | | | | | | | | | | 1.195*** | |
| | | | | | | | | | | (0.055) | |
| Econ*M′ | | | | | | | | | | -0.035 | |
| | | | | | | | | | | (0.087) | |
| M′=(-50, 0, 50) | | | | | | | | | | | 1.147*** |
| | | | | | | | | | | | (0.094) |
| Econ*M′ | | | | | | | | | | | -0.403*** |
| | | | | | | | | | | | (0.145) |
| Common Support | X | X | X | X | X | X | X | X | X | X | X |
| Constant | 0.223*** | 0.257*** | 16.978*** | 15.522*** | 24.192*** | 192.252*** | 186.535*** | 127.711*** | 100.156*** | 100.186*** | 98.064*** |
| | (0.020) | (0.020) | (0.261) | (0.216) | (0.268) | (7.912) | (8.190) | (3.742) | (4.301) | (2.241) | (3.790) |
| Observations | 849 | 849 | 849 | 849 | 849 | 294 | 289 | 430 | 419 | 848 | 848 |
| R-squared | 0.026 | 0.006 | 0.014 | 0.002 | 0.013 | 0.002 | 0.005 | 0.010 | 0.005 | 0.477 | 0.186 |

In this table, we limit the regression sample using a new criteria: students lying in the 0.15 times the standard deviation within the distribution of the CEE score.

The dependent variables in columns (1) and (2) are the share of risk neutral students in MPL 1 and MPL 2, in which we pool WTA and WTP together. Columns (3), (4), and (5) report results on probabilistic belief questions on the law of large numbers (LLN), two identical choices, and Exact Representativeness (ER), respectively. The outcome variables in columns (6) and (7) are the Dictators' actual sharing and Bystander's beliefs regarding the Dictators' sharing in the Dictator Game. Column (8) analyzes how an economics education affects students' sharing behavior as the Proposer in the Trust Game, which could be interpreted as students' beliefs regarding the amount that the other players would like to reciprocate. The dependent variable is Bystanders' belief regarding the mean amount of Player A's sharing in the Trust Game in column (9). Columns (10) and (11) ask Player B the amount she would like to give back if Player A gives 50, 100, 150 Yuan and Bystander's belief regarding the mean amount of Player B's giving back.

All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics major category (*Major-Preference FX*), and limit the regression sample to students who put both economics and non-economics majors in their rank-order list (*Common Support of Major Preference*).

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table B.9: Economics &Business Majors before Non-economics Majors

| | Risk Preferences | | Probabilistic Beliefs (3)-(5) | | | Dictator Game | | Trust Game (8)-(11) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| | MPL1 | MPL2 | Law of Large | Two Indifferent | Exact | Dictator's | Bystander's | Proposer's | Bystander's | Reciprocity of | Bystander's Belief |
| Dep. Var. | Risk Neutral | Risk Neutral | Numbers | Choices | Representiveness | Sharing | Belief | Sharing | Belief | Player 2 | about Reciprocity |
| Econ*M' | | | | | | | | | | -0.062 | -0.238* |
| | | | | | | | | | | (0.123) | (0.127) |
| Econ=1 | 0.147*** | -0.017 | 1.581*** | 0.464 | 1.727*** | -11.908 | -32.883* | -8.611 | 9.105 | -1.666 | -5.206 |
| | (0.043) | (0.045) | (0.610) | (0.505) | (0.593) | (17.252) | (18.525) | (8.467) | (10.317) | (5.024) | (5.188) |
| M'=(-50, 0, 50) | | | | | | | | | | 1.171*** | 1.221*** |
| | | | | | | | | | | (0.071) | (0.073) |
| Common Support | X | X | X | X | X | X | X | X | X | X | X |
| Constant | 0.186*** | 0.261*** | 16.634*** | 15.092*** | 24.088*** | 206.486*** | 184.966*** | 123.103*** | 94.943*** | 97.684*** | 97.356*** |
| | (0.024) | (0.025) | (0.342) | (0.283) | (0.332) | (9.387) | (11.005) | (4.808) | (5.703) | (2.817) | (2.907) |
| Observations | 430 | 430 | 430 | 430 | 430 | 152 | 136 | 214 | 216 | 430 | 430 |
| R-squared | 0.026 | 0.000 | 0.015 | 0.002 | 0.019 | 0.003 | 0.023 | 0.005 | 0.004 | 0.481 | 0.467 |

In this table, we limit the regression sample to students who always put economics &business majors before all the non-economics majors.
The dependent variables in columns (1) and (2) are the share of risk neutral students in MPL 1 and MPL 2, in which we pool WTA and WTP together. Columns (3), (4), and (5) report results on probabilistic belief questions on the law of large numbers (LLN), two identical choices, and Exact Representativeness (ER), respectively. The outcome variables in columns (6) and (7) are the Dictators' actual sharing and Bystander's beliefs regarding the Dictators' sharing in the Dictator Game. Column (8) analyzes how an economics education affects students' sharing behavior as the Proposer in the Trust Game, which could be interpreted as students' beliefs regarding the amount that the other players would like to reciprocate. The dependent variable is Bystanders' belief regarding the mean amount of Player A's sharing in the Trust Game in column (9). Columns (10) and (11) ask Player B the amount she would like to give back if Player A gives 50, 100, 150 Yuan and Bystander's belief regarding the mean amount of Player B's giving back.
All columns control for a vector of dummies that denotes whether students' majors in their rank-order list belong to the economics major category (*Major-Preference FX*), and limit the regression sample to students who put both economics and non-economics majors in their rank-order list (*Common Support of Major Preference*).
Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

## B.2 Details of Survey Design

### Risk Preferences Elicitation

There are three sets of questions to elicit students' risk preferences.

The first set of questions elicit students' either Willingness to Pay (WTP) or Willingness to Accept (WTA) between a given lottery and a series of monotonically increasing certain payoffs, {25 RMB, 30, 35,..., 55, 60}. The lottery pays 30 RMB with probability 0.25 and 60 RMB with probability 0.75, and is the same for all students. The students are, however, randomized into two groups where one group is asked about their WTP for this lottery and the other group is asked about their WTA for this lottery. Both WTP and WTA are elicited using a multiple price-list style table.

In the case of WTA, for each row, subjects are presented the money they would get had they sold out the lottery. Then, subjects are given two options, "sell" and "not sell", which are presented on the right side of the "selling price". The table below shows the structure of WTA mode:

Table B.10: WTA

| # of question | Price | Options | |
|---|---|---|---|
| 1 | 25 | Sell | Keep |
| 2 | 30 | Sell | Keep |
| 3 | 35 | Sell | Keep |
| 4 | 40 | Sell | Keep |
| 5 | 45 | Sell | Keep |
| 6 | 50 | Sell | Keep |
| 7 | 55 | Sell | Keep |
| 8 | 60 | Sell | Keep |

In the case of WTP, each subject is endowed with 60 RMB. For each row, subjects are presented the money they would have to pay had they bought the lottery. Then, subjects are given two options, "buy" and "not buy", which are presented on the right side of the "buying price". The table below shows the structure of WTP mode:

Table B.11: WTP

| # of question | Price | Options | |
|---|---|---|---|
| 1 | 25 | Buy | Not Buy |
| 2 | 30 | Buy | Not Buy |
| 3 | 35 | Buy | Not Buy |
| 4 | 40 | Buy | Not Buy |
| 5 | 45 | Buy | Not Buy |
| 6 | 50 | Buy | Not Buy |
| 7 | 55 | Buy | Not Buy |
| 8 | 60 | buy | Not buy |

As we can see from the preceding table, for the WTP mode, as the price increases from top to bottom, the deal becomes less and less appealing. Therefore, we expect a student who pays enough attention to such questions to select "Buy" first and then, at some point opt into "Not Buy" (of course, she could just choose "Buy" all the way from question 1 to 8). Our data suggests that this is the case: the vast majority of students answer in a consistent way with at most one switching point. The question where the answer differs from the previous question is called the "switching point".

For the second set of questions, we follow the price-list methodology developed by Holt and Laury (2002). Each decision row is a choice between Option A and B. Option A receives 30 RMB with probability $Pr = 0.25$, and 60 with $Pr = 0.75$. The series of Option B receives 400 with increasing probabilities $Pr(400) = \{0.01, 0.03, 0.05, 0.07, ..., 0.23, 0.25\}$, or receives nothing, with probability $1 - Pr(400)$. The structure of the elicitation is as follows:

Table B.12: The Second Set of Questions

| # of question | Option A | Option B |
|---|---|---|
| 9 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.01, pay 0 w.p.0.99 |
| 10 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.03, pay 0 w.p.0.97 |
| 11 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.05, pay 0 w.p.0.95 |
| 12 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.07, pay 0 w.p.0.93 |
| 13 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.09, pay 0 w.p.0.91 |
| 14 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.11, pay 0 w.p.0.89 |
| 15 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.13, pay 0 w.p.0.87 |
| 16 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.15, pay 0 w.p.0.85 |
| 17 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.17, pay 0 w.p.0.83 |
| 18 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.19, pay 0 w.p.0.81 |
| 19 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.21, pay 0 w.p.0.79 |
| 20 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.23, pay 0 w.p.0.77 |
| 21 | Pay 30 w.p.0.25, pay 60 w.p.0.75 | Pay 400 w.p.0.25, pay 0 w.p.0.75 |

This set of questions is similar to the previous. The left column presents a fixed option and the right column becomes better and better as the # of the question increases. We would expect same pattern (switching between options once at most), and, indeed, what we find verifies our expectation.

As Holt and Laury (2002) notes, such a manner of elicitation can characterize subjects' preferences because of the monotonicity of the column. Since one of the two options is fixed while the other becomes better and better (or worse and worse) over time, the presence of the switching point indicates that subjects' preferences change between the two particular questions around the switching point. Assume that there is an indifferent point such that the lottery in the column with varying lotteries brings about equivalent utility to the other column with a fixed lottery, then the switching point effectively bounds the indifferent point. This indifferent point is indicative of subjects' risk attitudes in this particular context, and we have discussed in Section 2.4 about how to relate the risk preference parameters to the switching point.

## Probabilistic Beliefs Elicitation

There are, in total, three questions for this part. All the questions are taking the form of an asset allocation problem: Students are asked to allocate their resources ( 30 virtual coins for each question) between two Arrow-Debreu assets, A and B. Asset A pays off if and only if event A is realized. By the same token, asset B pays off if and only if event B is realized. Each unit of asset A or B pays 1 lottery if event A or B is realized.

**Question 1**

Flip a fair coin 1,000 times.
Event A: the coin's head appears at least 530 times.
Event B (complements): the coin's head appears less than 530 times.
Allocate your resources on asset A and asset B such that asset A + asset B = 30.

**Question 2**

Flip a fair coin 10 times.
Event A: the coin's head appears in the ninth and tenth round. Event B: the coin's tail appears in the ninth and tenth round.
Allocate your resources on asset A and asset B such that asset A + asset B = 30.

**Question 3**

Flip a fair coin 100 times. Event A: the coin's head appears for exactly 50 rounds. Event B: the coin's head appears for either more or less than 50 rounds.
Allocate your resources on asset A and asset B such that asset A + asset B = 30.

## Social Preferences Elicitation

In the module of social preferences, students were asked to play a series of real-stakes games, wherein they received the payoff promised if their responses were randomly selected for reward.

Dictator Game: there are two players, A and B, in the game. In the first step, Player A receives 500 Yuan. She is then told to split the money between herself and the Player B. She can choose any amount she likes (from 0 to 500 Yuan) to keep, and give the rest to Player B. Player B can only accept what he gets from Player A. In terms of monetary incentive, Player A will get 500 minus the amount she/he sends out, and Player B will get the money that Player A is willing to transfer. In the Dictator Game, each participant is randomly assigned to one of the three scenarios: (a) if you are the Dictator (Player A), how much money out of 500 Yuan are you willing to share with Player B? (b) as a Bystander, what's your belief regarding the median value of the Dictators' sharing value in the Dictator Game? (c) you are the Receiver (Player B), no action is needed.

Ultimatum Bargaining Game: there are two players, A and B, in the game. In the first step, Player A receives 500 Yuan. She is then told to split the money between herself and Player B. She can choose any amount she likes (from 0 to 500 Yuan) to keep, and give the rest to Player B. In the second step, Player B can choose to accept, which results in the same outcome as the Dictator Game, or choose to decline, in which case both players get zero.

Trust Game: there are two players, A and B, in the game. In the first step, Player A could choose to send $X$ amount of 500 Yuan to Player B. Player A is also informed that what she sends would be tripled when Player B receives the money. In the second step, Player B gets three times the money from Player A (A is like an investor and B is like a manager). Player B is told to split the money he has between A and B: he can also choose whichever amount he likes (from 0 to all the money he has) to keep, and give the rest to Player A.

In the Trust Game, there are two anonymously paired players. Player A chooses to send amount $X$ out of 500 Yuan to Player B (including zero). Player A is also informed that whatever she sends will be tripled by our website. Consequently, when the Player A shares a value $X$ with Player B, the website will triple it, and give $3X$ to Player B. Player B then makes a similar decision —- gives some amount out of $3X$ back to Player A, including possibly zero. We constructed the question as "If you are Player B and know that Player A sends you $X$, the you will get $3X$. How much money are you willing to give back to Player A, when $X$ takes the values of {50, 100, 150}, respectively?" Additionally, we ask their beliefs as a Bystander regarding the median value of the amount given back, where $X$ takes the values of {50, 100, 150}.

Belief elicitation: We ask Bystanders to predict the median action of Player A in the Dictator Game and the Trust Game, and the action of Player B in the Trust Game following Krupka and Weber (2013).

The logistics are as follows. There are three roles in all games: Player A, Player B, and a Bystander. Every student is randomly assigned to one of the roles (so each of them will play one (potentially different) role in each of the games). Player A and B play what we describe previously; the Bystander elicits her beliefs regarding Player A and B's actions.

## B.3 Additional Results in Robustness Check and Heterogeneity of Treatment Effects

### Financial Status

Economics students in their fourth year may participate in more part-time internships and accordingly have differential financial situations. Luckily, the survey also includes questions about students' financial status, for example, the difference between income and spending in the survey month. We add variables on financial status together with all the other control variables to our main analysis. The results are summarized in Table B.6 (Appendix B1). In all columns, the coefficients for economics majors are highly consistent with our main results, implying that financial status is unlikely to drive our main results. We also test the robustness of our results by: (i) adding a variety of additional controls in the regression; (ii) altering the classification of treatment and control group. Results are reported in Table B.9 (Appendix B1).

### Major Preferences

As we have discussed in Section 2.2 and Section 2.3, conceptually we identify the causal effect by restricting our analysis to students who prefer economics major to other majors.

As we demonstrate in Section 2.2, a large majority of students indeed put all economics majors above other types of majors, suggesting that they unambiguously prefer economics majors. There are still, however some students whose ROLs are like what we show in Panel B of Table 2.1, where there are multiple switches between economics and non-economics majors. While these students still prefer an economics major to a non-economics major if we just look at the local major ranking, it becomes less clear whether these students always prefer an economics major unconditionally.

In this section we exclude the students aforementioned, namely, those whose ROLs exhibit multiple switch between economics and non-economics majors. The regressions are running based on equation (1) and (2). Again, the results reported in columns (1) -(11) of Table B.9 are quite robust and similar in significance and magnitudes to those results in Table 2.4, 2.6, 2.8, and 2.9.

## Success/failure in Major Application

An alternative explanation could be that a successful experience in major application (which leads to enhanced self-confidence), career validation, or a shock to future expected income, undermines our interpretation. This implies that the effect we observe should be strong even for the freshman and is homogeneous regardless of which compulsory courses taken, as these courses only teach students concepts common in academia, but not guidance of career development. Another possibility is that the impact on preferences is most salient in the first year and wanes as time goes by. However, we have demonstrated in previous sections that the magnitude of effect does depend on length of enrollment and curriculum for a variety of the outcome variables, and that the effect is fostered, rather than diminished, over time. Therefore, our results cannot be easily explained by immediate impacts brought about by success/failure in major application.

We also conduct additional analyses that test the effect of disappointment. If the disappointment alone explains the pattern in our data, we would expect students with higher pre-college rankings to be more disappointed than those with lower pre-college rankings, given that both are assigned to non-preferred majors. The results are presented in Table B.4. The first column reports the performance of top-ranking students who are assigned to non-economics majors, compared to the average of economics students. The second and third show the performance of middle-ranking and bottom-ranking students. It appears that the effects of pre-college ranking on all outcomes do not significantly differ, implying that disappointment is not the main driving force, otherwise the coefficients of the top-ranking non-economics students would be significantly larger than that of the middle ranking in absolute value.

Another piece of evidence that contradicts the disappointment effects is that among students who are not assigned to their most preferred majors, the position of students' admitted major in their rank-order list does not have a significant effect on decision-making variables. If the disappointment effect drives our main result, we would expect students with a lower admitted position to feel more unsatisfactory than those with a higher admitted position. Thus we estimate the following equation:

$$Y = \lambda_0 + \sum_{i=2}^{7} \lambda_i position_i + \omega econ + MajorFE + \epsilon. \tag{B.1}$$

Where $Y$ is subjects' response. $position_i$ is a vector of dummies that indicate the place of students' admitted major in their rank-order list. Therefore, $\lambda_i$ represents the effect of students' position of admitted major in their rank-order list on risk, social preferences. We also add an indicator of the economics & business major and students' major-preference fixed effects. If the confounding of admission position is at play, we would expect that $\lambda_i$ should be jointly significant. In the regression, we drop students who were admitted in the first position because only students who declared economics & business majors as their firstly preferred majors are included in our sample. We summarize the results of our regression analysis in Table B.5 (Appendix B1). In this table, Admitted in 3rd - Admitted in 6th denote that students were admitted in the third,...,sixth position of their rank-order list. And Admitted in 6th plus indicates that students didn't meet the cutoffs of the six major preferences, but were assigned to a major by CUFE. We conduct the F-test for the joint significance of the dummies for the position of admitted major in the rank-order list. The test results are reported at the middle of each column. The p-values of the test indicate that disappointment from an undesirable major is unlikely to affect our results.

## Heterogeneous Treatment Effects by Gender

Many studies have highlighted that higher education and academic economics have unequal treatment effects on students by gender. For example, women are less likely to receive credits from a co-authorship and get a promotion (Sarsons, 2017; Card et al., 2020); professor gender has a larger impact on female students' performance in science classes and their future development Carrell, Page, and West (2010). In this paper, we exploit a similar procedure to examine the relationship between economics education and gender difference. Table B.7 demonstrates the unequal (equal) treatment effects of economics education on the same outcomes as shown in Table B.4 using the following equation.

$$Y = \kappa' + \beta_1' econ + \beta_2' econ{*}male + \theta' X + \epsilon'''. \tag{B.2}$$

Where $Y$ is the subjects' response. We introduce the interaction term $econ*male$, which implies whether a participant is a male economics student or not. The coefficients on the interaction term $econ * male$ show the difference of outcomes between male economics students and female economics students. Only column (1) and (3) show week evidence on the unequal treatment effects of economics education: the treatment effects mainly concentrate on male students comparing to female students. Columns (4)-(7) indicate that, on average, male economics students show no difference relative to non-economics students in the indifferent-choice question, the probabilistic belief questions on exact representiveness and social preferences.

# Appendix C

# Chapter 3 Supplementary Tables and Figures

Table C.1: Balance Check for Household Size Increase Group

| (1)<br>Variable Name | (2)<br>Unmatched (C1) | (3)<br>Matched (T1) | (4)<br>Difference | (5)<br>SE | (6)<br>P-value |
|---|---|---|---|---|---|
| Household size | 3.973 | 3.913 | 0.060 | 0.086 | 0.491 |
| Rural members | 3.885 | 3.831 | 0.053 | 0.088 | 0.544 |
| Total land size (mu) | 7.678 | 8.382 | -0.704 | 0.962 | 0.464 |
| Rent-in land size (mu) | 0.370 | 0.190 | 0.180 | 0.122 | 0.139 |
| Total crop production (kg) | 2747.607 | 2820.620 | -73.013 | 215.539 | 0.735 |
| Forest land size (mu) | 2.316 | 2.552 | -0.236 | 0.801 | 0.768 |
| Fixed investment | 4575.034 | 3742.584 | 832.450 | 1123.082 | 0.459 |
| Ag, forestry, fishery machine | 361.850 | 592.664 | -230.814 | 203.922 | 0.258 |
| Industrial machinery | 214.848 | 317.113 | -102.266 | 143.301 | 0.476 |
| Transport machinery | 1302.509 | 999.258 | 303.251 | 517.239 | 0.558 |
| Expenditure on food crop seeds | 117.714 | 127.852 | -10.138 | 20.944 | 0.628 |
| Expenditure on food cash seeds | 12.653 | 12.641 | 0.012 | 2.651 | 0.996 |
| Expenditure on fertilizer | 775.237 | 744.161 | 31.076 | 77.412 | 0.688 |
| Total Income | 16040.612 | 15559.909 | 480.703 | 1727.880 | 0.781 |
| Total operation income | 11822.990 | 11131.765 | 691.226 | 1657.790 | 0.677 |
| Labor working off-farm | 0.284 | 0.278 | 0.006 | 0.042 | 0.894 |
| Migrating labor income | 1756.440 | 1787.829 | -31.389 | 274.105 | 0.909 |
| Total expenditure | 13957.491 | 13179.036 | 778.456 | 1680.843 | 0.643 |
| Total operation expenditure | 5386.776 | 5212.971 | 173.805 | 1246.511 | 0.889 |
| Total living expenditure | 7283.258 | 6486.478 | 796.780 | 846.488 | 0.347 |
| Annual borrow-in | 1522.542 | 1381.756 | 140.786 | 624.837 | 0.822 |
| Total labor input (days) | 446.806 | 445.651 | 1.155 | 23.628 | 0.961 |
| Ag labor input (days) | 241.538 | 246.376 | -4.837 | 12.498 | 0.699 |
| Off-farm labor input (days) | 75.286 | 70.078 | 5.208 | 12.609 | 0.680 |

Table C.2: Balance Check for Household Size Decrease Group

| (1) Variable Name | (2) Unmatched ((T2)) | (3) Matched (C2) | (4) Difference | (5) SE | (6) P-value |
|---|---|---|---|---|---|
| Household size | 4.510 | 4.570 | -0.059 | 0.076 | 0.436 |
| Rural members | 4.359 | 4.430 | -0.072 | 0.082 | 0.381 |
| Total land size (mu) | 7.891 | 8.514 | -0.622 | 0.558 | 0.265 |
| Rent-in land size (mu) | 0.217 | 0.221 | -0.005 | 0.081 | 0.952 |
| Total crop production (kg) | 2799.318 | 2974.005 | -174.687 | 169.958 | 0.304 |
| Forest land size (mu) | 1.859 | 2.883 | -1.024 | 0.716 | 0.153 |
| Fixed investment | 3873.866 | 3546.288 | 327.578 | 615.293 | 0.595 |
| Ag, forestry, fishery machine | 458.426 | 513.351 | -54.925 | 126.515 | 0.664 |
| Industrial machinery | 422.373 | 292.747 | 129.626 | 130.340 | 0.320 |
| Transport machinery | 1303.234 | 1060.653 | 242.581 | 528.343 | 0.646 |
| Expenditure on food crop seeds | 108.918 | 120.347 | -11.428 | 14.088 | 0.417 |
| Expenditure on cash crop seeds | 13.532 | 11.744 | 1.788 | 2.181 | 0.412 |
| Expenditure on fertilizer | 769.314 | 758.164 | 11.150 | 49.827 | 0.823 |
| Total income | 16325.852 | 14850.103 | 1475.749 | 1571.135 | 0.348 |
| Total operation income | 12087.470 | 10549.750 | 1537.720 | 1471.565 | 0.296 |
| Labor working off-farm | 0.268 | 0.309 | -0.041 | 0.036 | 0.254 |
| Migrating labor income | 1496.418 | 1710.437 | -214.019 | 219.719 | 0.330 |
| Total expenditure | 14528.858 | 13142.854 | 1386.004 | 1440.737 | 0.336 |
| Total operation expenditure | 5078.027 | 4526.813 | 551.214 | 1159.289 | 0.635 |
| Total living expenditure | 7716.856 | 6944.946 | 771.910 | 484.319 | 0.111 |
| Annual borrow-in | 1850.979 | 1624.378 | 226.601 | 456.167 | 0.619 |
| Total labor input (days) | 453.482 | 456.702 | -3.220 | 16.124 | 0.842 |
| Ag labor input (days) | 249.782 | 263.113 | -13.331 | 10.428 | 0.201 |
| Off-farm labor input (days) | 67.611 | 68.815 | -1.204 | 8.883 | 0.892 |

Table C.3: Household Land Size Change in Response to Reallocation

| | HH Increase | | HH Decrease | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Asinh Land Per Capita | Asinh Land Per Labor | Asinh Land Per Capita | Asinh Land Per Labor |
| **Panel A: include $T-4$ period** | | | | |
| Pop change before reallocation | 0.055*** | 0.077*** | -0.071*** | -0.042** |
| | (0.020) | (0.025) | (0.016) | (0.019) |
| Observations | 7788 | 7688 | 11185 | 10972 |
| **Panel B: include $T+5$ period** | | | | |
| Pop change before reallocation | 0.059*** | 0.069*** | -0.075*** | -0.042** |
| | (0.021) | (0.026) | (0.015) | (0.018) |
| Observations | 8038 | 7913 | 11581 | 11327 |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |

*Note:* This table serves as the robustness check by changing the definition of our treatment and control groups. It demonstrates that how does household land size vary by the family size change before and after the last reallocation. Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table C.4: Household Labor Change in Response to Reallocation

| | HH Increase | | HH Decrease | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Agricultural Labor** | Ag Labor Input(days) | Ag Labor Input Per Mu | Ag Labor Input(days) | Ag Labor Input Per Mu |
| **Panel A: include $T-4$ period** | | | | |
| Pop change before reallocation | -0.065 | -0.083* | -0.145** | -0.025 |
| | (0.055) | (0.045) | (0.061) | (0.044) |
| Observations | 7788 | 7526 | 11185 | 10719 |
| **Panel B: include $T+5$ period** | | | | |
| Pop change before reallocation | -0.089 | -0.098** | -0.160*** | -0.045 |
| | (0.058) | (0.048) | (0.061) | (0.046) |
| Observations | 8039 | 7773 | 11583 | 11092 |
| **Non-agricultural Labor** | Total Nonag | Outside Village | Total Nonag | Outside Village |
| **Panel C: include $T-4$ period** | | | | |
| Pop change before reallocation | -0.331** | 0.072 | -0.144 | -0.084 |
| | (0.129) | (0.109) | (0.125) | (0.128) |
| Observations | 5842 | 6248 | 8966 | 9357 |
| **Panel D: include $T+5$ period** | | | | |
| Pop change before reallocation | -0.309** | 0.026 | -0.145 | -0.044 |
| | (0.129) | (0.110) | (0.121) | (0.128) |
| Observations | 5820 | 6005 | 8866 | 9012 |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |

*Note:* This table serves as the robustness check by changing the definition of our treatment and control groups. It demonstrates that how does household labor input change in response to land reallocation.
Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table C.5: Household Land and Labor Change in Response to Reallocation

| | HH Increase | | HH Decrease | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A: Land Change** | | | | |
| | Asinh Land Per Capita | Asinh Land Per Labor | Asinh Land Per Capita | Asinh Land Per Labor |
| Pop change before reallocation | 0.083*** | 0.062** | -0.122*** | -0.069** |
| | (0.023) | (0.028) | (0.024) | (0.030) |
| Observations | 3354 | 3299 | 4486 | 4398 |
| **Panel B: Agricultural Labor Input** | | | | |
| | Ag Labor Input(days) | Ag Labor Input Per Mu | Ag Labor Input(days) | Ag Labor Input Per Mu |
| Pop change before reallocation | -0.117 | -0.080* | -0.256*** | -0.103 |
| | (0.074) | (0.045) | (0.092) | (0.066) |
| Observations | 3354 | 3234 | 4486 | 4340 |
| **Panel C: Non-agricultural Labor Input** | | | | |
| | Total Nonag | Outside Village | Total Nonag | Outside Village |
| Pop change before reallocation | -0.412** | -0.045 | 0.060 | 0.060 |
| | (0.193) | (0.173) | (0.178) | (0.177) |
| Observations | 2448 | 2601 | 3550 | 3710 |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |

*Note:* This table serves as the robustness check by changing the definition of our treatment and control groups. It demonstrates that how does the timing of household family size change affect household land size/labor input.
Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table C.6: Household Land Size Change due to Land Reform

| | HH Increase (+) | | HH Decrease (-) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Asinh | Asinh | Asinh | Asinh |
| | Per Capita | Per labor | Per Capita | Per labor |
| Reform × | 0.060** | 0.064** | -0.067** | -0.054 |
| Pop change after | (0.027) | (0.023) | (0.027) | (0.039) |
| Reform × | 0.012 | 0.005 | 0.014 | 0.014 |
| Pop change before | (0.023) | (0.032) | (0.032) | (0.030) |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 2450 | 2330 | 3335 | 3166 |

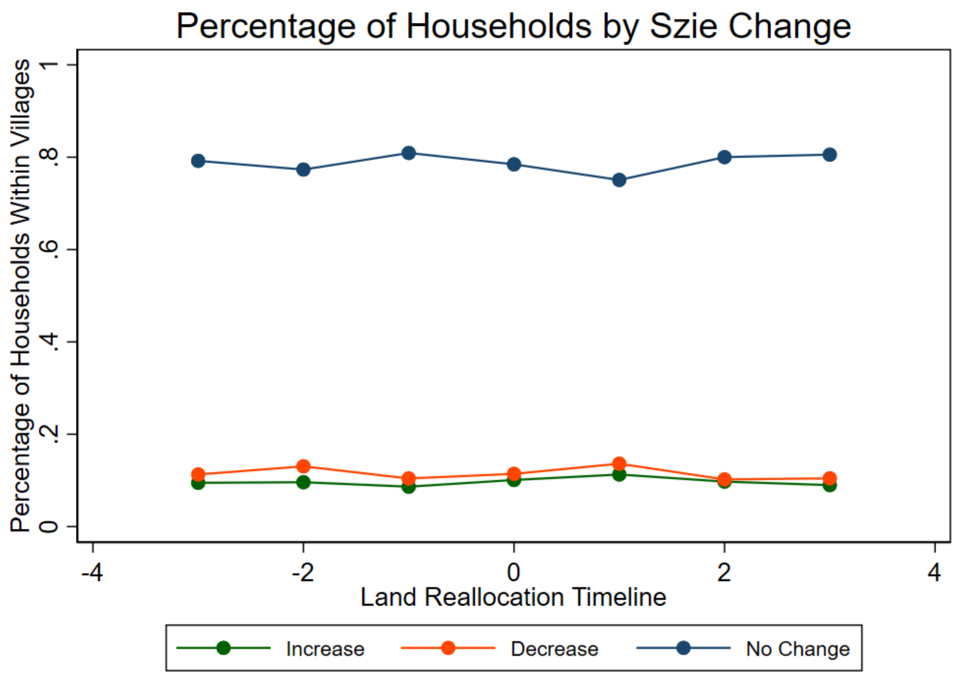
*Note:* This table serves as the robustness check by changing the definition of our treatment and control groups. It demonstrates that how does the land market reform affect household labor input.
Robust standard errors clustered at the village level are reported in brackets. *** indicates 1% significance; ** 5%; and * 10%.

Table C.7: Household Labor Change due to Labor Reform

| | HH Increase (+) | | HH Decrease (-) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Ag labor | Migration Inc | Ag labor | Migration Inc |
| Reform × | -0.927** | -0.727** | -0.947*** | -0.660*** |
| Pop change after | (0.395) | (0.279) | (0.280) | (0.220) |
| Reform × | 0.033 | -0.024 | -0.541* | -0.394** |
| Pop change before | (0.251) | (0.232) | (0.290) | (0.197) |
| Household FE | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 1789 | 1767 | 2496 | 2449 |

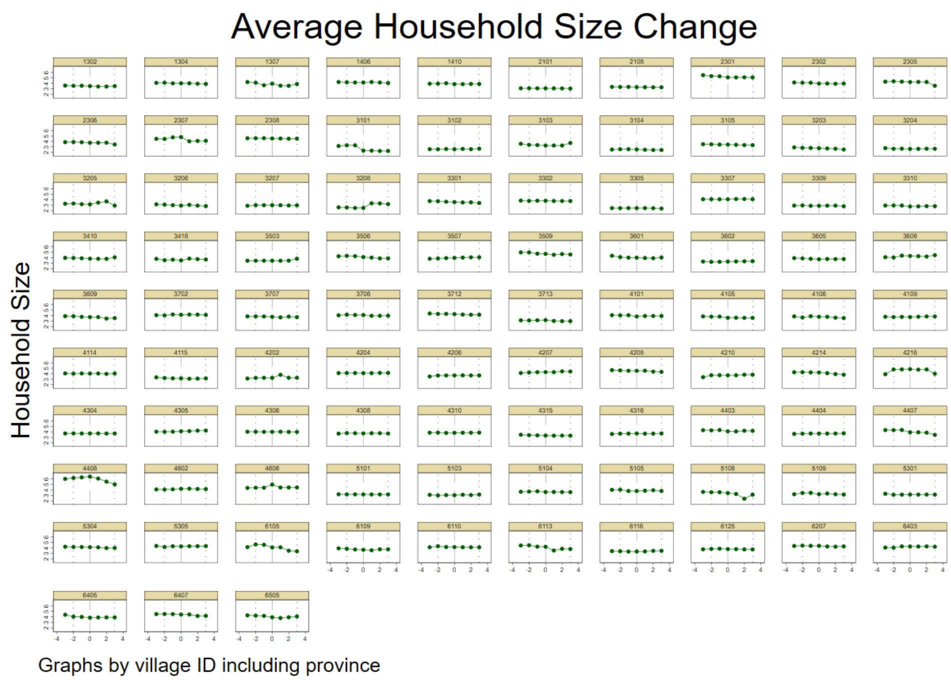
*Note:* This table serves as the robustness check by changing the definition of our treatment and control groups. It demonstrates that how does the labor market reform affect household labor input.
Robust standard errors clustered at the village level are reported in brackets.
*** indicates 1% significance; ** 5%; and * 10%.

Figure C.1: Share of Households by Size Change and Land Redistribution Timeline




Note: This figure presents the share of households within that village experience different changes in family size across the timing of land reallocation.

Figure C.2: Demographic Dynamics at the Village Level




Note: This figure presents the pattern of households family size at the village level across the timing of land reallocation.