Nicholas Burger
Mary Fu
Kun Gu
Xiangping Jia
Krishna B Kumar
Guo Mingliang

Assessing the impact of farmer field schools on fertiliser use in China

March 2015

Impact Evaluation Report 25

Agriculture

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programmes. We are the global leader in funding and producing high-quality evidence of what works, how, why and at what cost. We believe that better and policy-relevant evidence will make development more effective and improve people's lives.

3ie Impact Evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of this report as partial fulfilment of requirements under grant OW3.1216 issued under Open Window 3. The content has been copyedited and formatted for publication by 3ie. Due to unavoidable constraints at the time of publication, a few of the tables or figures may be less than optimal. All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its Board of Commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Krishna B Kumar at kumar@rand.org

Funding for this impact evaluation was provided by 3ie's donors, which include UKaid, the Bill & Melinda Gates Foundation, Hewlett Foundation and 12 other 3ie members that provide institutional support. A complete listing is provided on the 3ie website at http://www.3ieimpact.org/about-us/3ie-members/

Suggested citation: Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G, 2015. *Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25.* New Delhi: International Initiative for Impact Evaluation (3ie)

3ie Impact Evaluation Report Series executive editors: Jyotsna Puri and Beryl Leach
Managing editor: Omita Goyal
Assistant managing editor: Kanika Jha
Production manager: Pradeep Singh
Copy editor: Arpita Das
Proofreader: Aruna Ramachandran
Cover design: John F McGill
Printer: VIA Interactive
Cover photo: International Rice Research Institute

Assessing the impact of farmer field schools on fertiliser use in China

Nicholas Burger
RAND Corporation

Mary Fu
Pardee RAND Graduate School

Kun Gu
Pardee RAND Graduate School

Xiangping Jia
Chinese Center for Agricultural Policy

Krishna B Kumar
RAND Corporation

Guo Mingliang
Chinese Center for Agricultural Policy

3ie Impact Evaluation Report 25
March 2015

International Initiative
for Impact Evaluation

Acknowledgements

We thank 3ie and its anonymous reviewers for feedback that has considerably improved this report. We are also grateful to Dr Jikun Huang, CCAP, for his tremendous technical advice and organizational support, and Dr Fusuo Zhang, CAU, for his many insights into Chinese agriculture. Thanks to our research assistant, Ma Tao, for excellent research and field support. Gu and Kumar acknowledge the support of the Rosenfeld Program on Asian Development at the Pardee RAND Graduate School.

Abstract

In China, a major agricultural challenge is the sub-optimal use of fertilizer and the environmental effects associated with overuse. The Chinese Ministry of Agriculture (MoA) is addressing this problem by instituting farmer field schools (FFS), but this initiative has not been rigorously evaluated. We used a randomised controlled trial to evaluate the FFS program in five counties in Anhui and Hebei provinces, for rice and tomato crops, respectively. We used a matched pair random assignment of villages into treatment and control groups, and we randomized additional farmers into an 'exposed' group to study diffusion effects. We found no significant effects of the FFS intervention on mean fertilizer use for either crop. However, we found that fertilizer usage is highly heterogeneous, and a simple comparison of means masks the differential response to the FFS programme at either end of the distribution. For rice farmers, the percentage increase in nitrogen fertilizer usage at the lowest quintile is significantly higher for the treatment group than that for the control group, with a less pronounced drop in usage in the highest quintile. Ordinary least squares and instrumental variables regressions confirm that the distance from the prescribed optimum fertilizer use for rice decreased due to the intervention. For tomato farmers, nitrogen use increased in the lowest quintile more in the treatment group than in the control group, but the reduction in the highest quintile in the control group was substantially higher than that in the treatment group. Overall we conclude that the FFS programme improved the optimal use of fertilizer for rice farmers but had insignificant effects for tomato growers. Given the inconclusive results, we conclude that policymakers should revisit plans to scale up FFS in China, paying special attention to crop specificity, heterogeneous implementation quality, and outcomes not limited to fertilizer usage.

Contents

List of figures and tables

Abbreviations and acronyms

CCAP	Center for Chinese Agricultural Policy
FFS	farmer field schools
GHG	greenhouse gas
GHV	greenhouse vegetable
K	Potassium
MoA	Ministry of Agriculture
N	Nitrogen
P	Phosphorus
RCT	randomized controlled trial
TOT	training of trainers
OLS	ordinary least squares
IV	instrumental variable

1. Introduction

In China, a major agricultural challenge is the inefficient use of fertilizer and the environmental effects associated with its overuse. China's farmers use more fertilizer per hectare (more than 200 kg/ha) than farmers anywhere else in the world except for Japan, the Netherlands and South Korea. Existing studies have shown that overuse of nitrogen (N) fertilizer ranged from 30 per cent to 50 per cent in grain production (Huang *et al*. 2008). This excessive use has resulted in serious food safety and environmental problems, such as large N losses through NH3 volatilization and nitrogen-leaching into ground water, rivers and lakes (Xing and Zhu 2000; Zhu and Chen 2002). Because 70 per cent of agricultural greenhouse gas (GHG) emissions originate from N fertilizers, improved N management is critical for income of farmers and addressing climate change.

Our research evaluates a promising mechanism to address inefficient fertilizer use in China, the farmer field school (FFS). Research suggests that insufficient farmer knowledge and information about the effects of excess fertilizer is one reason for inefficient rates of nitrogen fertilizer application in China (Huang *et al*. 2008). However, given the large heterogeneity in fertilizer usage, it is unclear whether all farmers are using fertilizer in excess or whether the problem is one of farmers not applying fertilizers optimally. Moreover, the lack of accountability has made China's current public agricultural extension system ineffective at delivering fertilizer training and knowledge to individual farmers (Hu *et al*. 2009). The Chinese Ministry of Agriculture (MoA) is addressing this problem by instituting farmer field schools, hoping to avoid the pitfalls of the traditional extension system by using local farmer-trainers to improve accountability and effectiveness through a participatory approach to agricultural extension. However, a rigorous evaluation of China's FFS has not been conducted to date, and this is the gap we seek to fill with this evaluation. By evaluating the effectiveness of FFS, we may also be able to assess the potential for scaling up FFS in a cost-effective way in China.

While the intervention is based in China, the findings of our study might have implications beyond China. For instance, recent reports suggested that overuse of fertilizers is a problem in India as well.[1] Since China and India are the two most populous countries with large shares of agricultural labour force, and have had similar experiences with the Green Revolution, any study that sheds light on improving farming decisions in these countries could have far-reaching implications.

The overall goal of this project is to evaluate the impact of fertilizer-related training provided by FFS to Chinese farmers. The following questions are of particular interest:

- Do FFS graduates apply N fertilizers and other agro-chemical inputs more optimally?

[1] 'Green Revolution in India wilts as subsidies backfire'. *Wall Street Journal* [online], 22 February 2010. Available at: <http://online.wsj.com/article/SB10001424052748703615904575052921612723844.html> [Accessed 11 March 2015]

- Does the FFS programme lead to improved perceptions of environmental problems related to excessive fertilizer usage?
- Is there any knowledge diffusion from trained farmers to other farmers?
- What are the socioeconomic impacts, e.g. impacts on farmers' incomes, farm management capability, and farmers' perception of and behavior toward local institutions such as FFS?
- How cost-effective are FFS?
- How do the above impacts differ between greenhouse and grain farmers? Should China use FFS as one of the primary extension tools for its agricultural extension system?
- Should FFS be included in China's national policy for climate change?

We address many of these questions in this report. We have collected rich data through our baseline and endline surveys, which can allow us and other researchers to pursue any remaining questions in detail.

Since the FFS programme was delivered at the village level, we designed and implemented a clustered randomized controlled trial (RCT). Working with the MoA, we chose Anhui and Hebei provinces to conduct the intervention for rice and tomatoes, given their sizable production of the respective crops. We focus on rice and tomato farming because they are rapidly growing sources of N2O emissions and methane (Hu *et al.* 2009). Moreover, tomatoes are a greenhouse vegetable (GHV) and have significantly different fertilizer needs than rice. By choosing two very different crops in two very different provinces, we hope to identify how the effectiveness of the FFS programme varies by crop and location.

We conducted detailed power calculations for each crop separately to determine the number of villages and farmers per village. In Anhui, we chose two counties, four townships in each county, and seven randomly selected villages from each township for a total of 56 villages. We randomized 28 villages into the treatment group (received FFS training) and 28 into the control group (did not receive FFS training). In treatment villages, we randomly selected 10 farmers to be 'exposed' farmers to study diffusion effects. Our total sample size for rice counties is 1,120 farmers. Our approach for tomato farmers mirrored the approach for rice farmers. We chose three counties and 36 villages, for a total of 720 farmers. The programmes were rolled out between April 2012 and May 2013, and we carried out baseline and endline surveys before and after the interventions.

We find some—but not conclusive—evidence that FFS changed farmer behaviour and improved outcomes. As we explain in greater detail later in the report, we focus not just on changes in mean fertilizer use, where we find no FFS effect, but on differential changes along the distribution of fertilizer application. Fertilizer use is heterogeneous, and while some farmers in our baseline sample are using excessive amounts of fertilizer, others use too little. There is evidence that the FFS increased use by rice and tomato farmers at the lower end of distribution; there is limited evidence, and only for rice farmers, that the programme reduced use among high users.

The rest of this report is organized as follows. In Section 2 we present the description of the intervention and the theory of change. In Section 3 we present the institutional context. In Section 4 we discuss the timeline, and in Section 5 the methodology and the experimental design. In Section 6 we discuss the challenges with the implementation of the intervention. Section 7 presents the results of the evaluation and Section 8 presents a cost-effectiveness analysis. We discuss policy implications and recommendations in Section 9. Further details on the experimental design and power calculation are presented in the appendixes.

2. Description of intervention, theory of change and research hypothesis

2.1 Intervention

Policymakers in China have begun to focus on the problem of inefficient, and, in particular, excessive fertilizer use by farmers and have introduced FFS to disseminate knowledge to farmers. We evaluated the FFS intervention, financed by China's MoA and administered through local governments. We evaluated the effectiveness of training in optimizing fertilizer usage by Chinese GHV (tomato) and grain (rice) farmers. FFS training at the village level includes hands-on, farmer-managed learning on experimental plots, along with informal training prior to a single crop-growing season. Through group interaction, the goal of the FFS is to empower FFS graduates with skills in crop management, learning capabilities and communication.[2] Working with the MoA, we selected one extension agent for every one or two villages. These extension agents were trained before the intervention on the unified course content. Throughout the crop season, they disseminated low carbon farming practices to the villager farmers who are in the treatment group through lectures, field experiments and interactive communication.

To provide effective training that is targeted at local needs and conditions, the FFS curriculum was designed based on soil tests and fieldwork conducted by agricultural experts before the intervention and experiment began. The MoA developed the curriculum through a multi-step process. First, the MoA organized an advisory meeting that included FFS experts, soil scientists, agronomists and extension experts to brainstorm over curriculum design. Second, based on their previous work in the study area, natural scientists gave recommendations on fertilizer use, pest management, irrigation and other farm practices. Third, the recommendations were calibrated by local experts at the provincial and county levels. Fourth, the 'package' of technologies was reviewed by the FFS experts to design standards and detailed implementation guidelines. During the training of trainers (TOT) workshop, both the technology packages and FFS guidelines were communicated to the extension agents selected for our study.

Our RCT focuses primarily, although not exclusively, on fertilizer use, and the FFS programme provides guidance to farmers on how and how much nitrogen, phosphorus, and potassium (NPK) fertilizer to apply.[3] In Anhui province, one of the main training goals for fertilizer was to adjust the amount of N fertilizer farmers apply to 165–180 kg/ha, which is considered optimal by agronomists for 'normal' weather. In other words, for farmers who apply fertilizer excessively, the goal was to reduce usage to 165–180 kg/ha, while for those who use less than the optimum, the FFS sought to increase their

[2] The details of the FFS curriculum are provided in the appendixes.

[3] We do not address phosphorus fertilizer use in our study, since it was not regarded as being applied sub-optimally in our pre-study development work.

fertilizer use to improve yields.[4] In addition, the FFS sought to increase K fertilizer use in Anhui to avoid 'lodging' disease (described in detail in Section 7). In Hebei province, the guidance for tomato growers included recommendations for organic fertilizer use (typically cow manure) and chemical (manufactured) fertilizer. Chemical and organic fertilizers have different effects on soil quality, environmental impacts and costs. We discuss differential application rates by farmers in Section 7, but we leave more detailed analysis of the specific impacts of each kind of fertilizer for future work. Tables 2.1 through 2.3 outline the highlights of the FFS curriculum.

Table 2.1. FFS curriculum: recommended technology guidance in Anhui (rice farming)

Technology	Content
Fertilizer use	Total amount of N fertilizer use should be 165–180 kg/ha Applying fertilizer in jointing-booting stage Increasing K fertilizer use to avoid lodging
Crop protection	Helping farmers to identify main plant diseases through participation of FFS（false smut, leaf blast, panicle rice blast, sheath blight, plant-hoppers, leaf-roller, rice stem borer） Teaching farmers commonly used control methods and integrated control measures Changing commonly held wrong conceptions and methods on fertilizer use Enhancing the environmental and ecological awareness of the farmers
Cultivation	Recommending anti-lodging varieties Improving and enhancing the transplanting density Drying paddy field in sunshine to ensure effective tillers Changing 'cutting down water supply in the late period' behaviour
Response to unusual weather	Early drought (adjusting seeding and transplanting time) High temperature damage in flowering period (delaying sowing date) Irrigation and drainage during typhoon period (and lasting rain period) Sheath blight and false smut caused by typhoon and lasting rain Pre- and post-low temperature period (selection of species)

[4] This optimum range is consistent with generally optimal fertilizer use recommended by agronomists. In their 2002 paper on nitrogen fertilizer use in China, Zhu and Chen note, 'Crop yield is governed by a series of factors, some of which are difficult to predict. Therefore, even if the optimum N application rate is a rough range, it is still much better than applying without guidance. From the data obtained in some long-term field experiments conducted on the major crops in agricultural regions, a general range of N application rate for cereal crops is recommended as 150–180 kg N ha−1. In practice, it should be adjusted according to the local conditions (such as variety, irrigation, etc.).'

Table 2.2. FFS curriculum: recommended technology guidance in Hebei (tomato farming, short growing season)

Technology	Content
Irrigation and fertilization	1) Control excessive application of organic fertilizers. Base fertilizer should be 30,000–45,000 kg/ha (if using cow manure, then it should be 45,000–60,000 kg/ha). Old vegetable plots that have been used over 5 years should apply 15,000–30,000 kg/ha of straw compost or bio-organic fertilizer. 2) Base chemical fertilizer should use 900–1,200 kg/ha of superphosphate. Phosphatic fertilizer should use 10–15% less for old vegetable plots. Do not use chemical N and K fertilizer. 3) The total amount of after fertilizer used during the whole growth period should be 300–375 kg/ha of N and 450–600 kg/ha of K2O. After fertilizer should reduce 15–20% every time for old vegetable plots. 4) After fertilizer should be applied based on the growth stage. The first fertilizer starts when the fruit is the size of a walnut. 4–5 times of after fertilizer is fine. 5) Strictly control the amount of irrigation to prevent excessive humidity, which may cause diseases.

Table 2.3. FFS curriculum: recommended technology guidance in Hebei (tomato farming, long growing season)

Technology	Content
Irrigation and fertilization	1) Control excessive application of organic fertilizers. Base fertilizer should be 45,000–60,000 kg/ha (if using cow manure, then it should be 60,000–75,000 kg/ha). Old vegetable plots that have been used over 5 years should apply 30,000–45,000 kg/ha of straw compost or bio-organic fertilizer. 2) Base chemical fertilizer should use 1,200–1,500 kg/ha of superphosphate. Phosphatic fertilizer should use 10–15% less for old vegetable plots. Do not use chemical N and K fertilizer. 3) The total amount of after fertilizer used during the whole growth period should be 525–600 kg/ha of N and 525–600 kg/ha of K2O. After fertilizer should reduce 15–20% every time for old vegetable plots. 4) After fertilizer should be applied based on the growth stage. The first fertilizer starts when the fruit is the size of a walnut. 6–8 times of after fertilizer is fine. 5) Strictly control the amount of irrigation to prevent excessive humidity, which may cause diseases.

2.2 Theory of change

Our theory of change is predicated on achieving optimal N and K fertilizer use brought about by improved human and social capital imparted through participatory FFS.

Analysis of the context

The agriculture system in China—where sub-optimal amounts of fertilizer are used—has been described in greater detail in Sections 1 and 3. Earlier attempts at using extension agents to educate farmers have not been successful in this context. The FFS programme, which involves participatory training, is being tried by the Chinese MoA to remedy the situation. The theory of change described in this section is relevant to this context.

Assumptions

The key assumptions underlying the theory of change for the FFS intervention are:

- Chinese farmers are using sub-optimal (both in a private and social sense) amounts of N and K fertilizers for a variety of reasons.
- The optimal amounts of fertilizer usage per hectare, conditional on weather, can be determined.
- Farmer field schools can incentivize extension agents to impart more effective training to farmers.
- Training imparted by the FFS can exert a positive influence on farmer decisions.
- The imparted knowledge and change in fertilizer usage can be measured.

Causal pathways

Our theory of change is depicted in Figure 2.1, including the inputs, outputs, outcomes (and the causal pathways between the outputs and outcomes) and impacts.

A variety of factors could be responsible for the sub-optimal use of fertilizers by Chinese farmers:

- Perhaps the most straightforward explanation of this phenomenon is that the farmers lack sufficient knowledge of the optimal amount of fertilizer to use. Since excessive use of fertilizer can lead to environmental degradation through leaching into water sources and creating greenhouse gases, the lack of knowledge is related to both private as well as socially optimal levels. Under-usage of fertilizer would have the greatest impact on private returns, while over-usage would have the greatest impact on social returns. Imparting knowledge is perhaps the most obvious channel by which FFS could influence outcomes.
- Farmers may be aware of the environmental impact of their actions, but might not act on that knowledge if they perceive that the effects are distant or diffused.
- Farmers could be following practices they have followed in the past, or handed down to them by older generations, without considering whether their behaviour is optimal. In other words, behavioural reasons (for instance, loss aversion) or habit persistence could be behind observed farmer behaviour.
- A related reason could be that farmers are risk-averse. Farmers might be afraid that if they applied less fertilizer, their harvest might be at risk.
- Farmers might lack the opportunities to interact with other farmers in a structured environment to learn best practices in fertilizer usage from each other

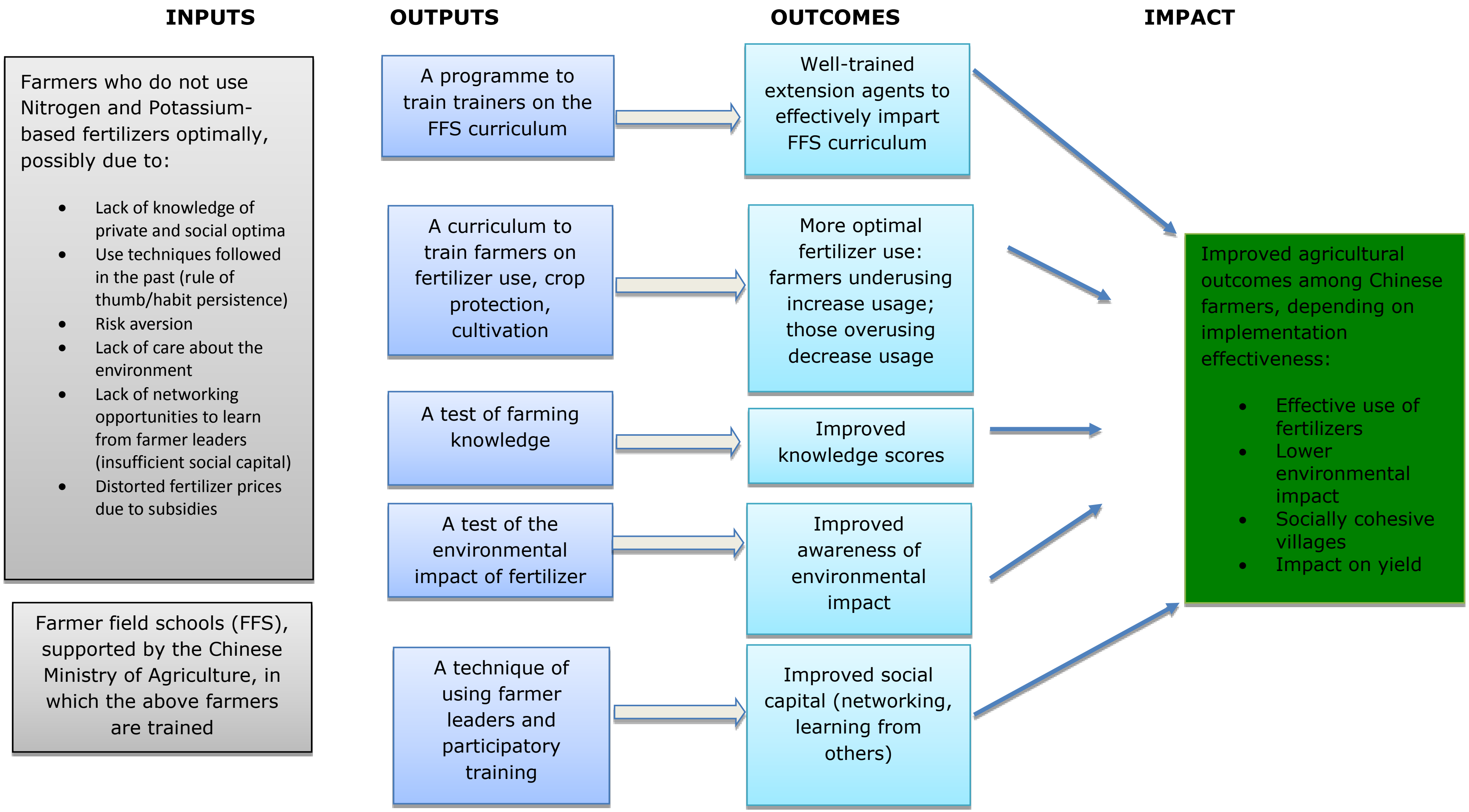
and farmer leaders. In other words, insufficient social as well as human capital might be behind the observed farmer behaviour.

- Subsidized prices could also cause farmers to use excessive fertilizer. While the FFS intervention could address the above reasons for sub-optimal fertilizer usage, it cannot address the problem of price distortions.

Attributing sub-optimal fertilizer usage to one of the reasons is not our primary focus. However, detailed analysis of the baseline data could presumably disentangle these various causes. For instance, we could check if farmers with a higher level of education use fertilizers more optimally; a higher level of experience might not yield such clear-cut results (more experience could mean more knowledge, but could also contribute to the persistence of bad practices). Instead our focus is the expected outcomes from hypothesized causes through which the FFS intervention could have an impact.

- The premise of the FFS is that using local farmer-trainers will improve accountability and effectiveness through a participatory approach to agricultural extension. However, the extension agents (trainers) would need to be trained themselves to teach the FFS curriculum. The outcome of such training is more effective extension agents. At this point, we have included extension agent fixed effects in a regression of impact and find that many of the agents show significant effects. In future, we will use data from an extension agent survey we conducted on age, education, experience, and so on, of extension agents to see what factors make a more effective extension agent.

Figure 2.1: Impact of FFS on fertilizer usage: theory of change



The core of the FFS curriculum imparts knowledge about the optimal use of fertilizer, protection of crops from diseases, better cultivation techniques, and altering fertilizer usage based on weather conditions. If the FFS is successful in imparting this knowledge, the outcome should be a move towards the optimum for fertilizer usage. Given the large amount of heterogeneity that exists in baseline fertilizer usage, the effect of training based on this curriculum should be increased usage at the lower end of the distribution and decreased usage at the upper end. We find strong evidence of this in treatment villages, especially for rice.

Improved knowledge of farming practices is a critical pathway that we hypothesize about in assessing the effectiveness of FFS. A good test of this pathway is improved knowledge scores in the test that was developed as part of the FFS curriculum, between the baseline and endline surveys. This is the result we find. Enhancing the environmental and ecological awareness of farmers was a key goal of FFS training. The effectiveness of the curriculum in bringing this about can be tested using the questions we included in the surveys.

Since the FFS by its very nature is participatory and based on a cadre of demonstration farmers (who are chosen for being key nodes in farmer networks), the FFS should improve social capital and the ability of farmers to learn from each other and transmit best practices.[5] We have collected information on the other key farmers each farmer interacts with, and using this we should be able to map the flow of influence across farmers in the village and the strengthening of these relationships between the baseline and the endline surveys. Improved social capital could result in better outcomes, presumably through increased information exchange, peer pressure and so on.

While the FFS curriculum and training programmes were designed to be uniform across villages, variability in implementation of the programme across villages is inevitable. This is the reason the impact box in Figure 2.1 notes the conditionality of impact on the effectiveness of implementation. In particular, since the main channel through which we expect FFS to have an effect is by increasing knowledge, we need to be aware of the institutional and other contexts that could impede the transmission of knowledge. For instance, if the FFS trainers are not trained or incentivized properly, effective training and transmission of knowledge is unlikely to occur. We have used village-level controls in our analysis to the extent possible to account for the variability. For tomatoes, we have dropped one county from our analysis where we suspect that the FFS training has not been implemented effectively.

[5] In rural communities, farmers are segmented into different groups. In the old socialist era, these were called production teams (*sheng chan dui*). Within each village, some households are selected into a committee that organizes village governance (including collecting tax previously; nowadays it means organizing public services and information collection and dissemination). In addition, many of the government programmes are coordinated through these farmers. For example, to introduce agricultural technologies, some of these farmers are selected as a 'demonstration base'.

2.3 Outcomes

The primary and secondary outcomes emerging from the theory of change analysis are summarized in Table 2.4. The categorization and listing of the outcomes should not be taken to imply that we can or will be able to test each of them. For instance, the environmental impact would be small for a pilot intervention and very difficult to detect, especially in the short term.

Table 2.4. Primary and secondary outcomes

Category	Outcome
Primary	• Optimal usage of fertilizer • Improved knowledge (scores) • Improved awareness of environmental impact of fertilizer usage • Lower environmental impact • Well-trained extension agent
Secondary	• Optimal pesticide usage • Lack of decrease in yield due to changes in fertilizer usage • Improved social capital

Beneficiary populations

The intended beneficiaries of the FFS intervention and evaluation are:

- **Farmers,** who were taught improved farming practices.
- **Extension agents**, who were trained initially on the FFS curriculum to become effective change agents.
- **Chinese agricultural policymakers** at the local and MoA levels, who can use the results from the evaluation to decide whether and how to scale up the FFS programme.
- **The Chinese public**, who would benefit from a sustainable and environmentally friendly way of using fertilizers.
- **Farmers and policymakers in other countries** who could use the results from the evaluation to design similar programmes to target excessive fertilizer use and other agriculture-related challenges.

3. Context

Previous studies have shown that the overuse of N fertilizer in China ranged from 30 per cent to 50 per cent in grain (e.g. rice) and vegetable (e.g. tomato) production, which has resulted in serious environmental problems. While there are a number of hypotheses for fertilizer overuse in China, ‘insufficient knowledge and information’ is believed to be the primary explanation. Huang *et al*. (2008) found that when farmers received training and in-the-field guidance, they were able to reduce N fertilizer use by as much as 35 per cent in rice production without lowering yield. Huang *et al.* (2010) found that maize farmers reduced N fertilizer use by 20 per cent with just two hours of training.

As in many countries, public extension services in China are the most common method of providing widespread information and training to farmers. The whole system consists of five levels: central government, province, city, county and town. Among them, the county-level and town-level extensions are the basic management and implementation units run by the local governments, providing service to farmers directly. From the perspective of specialized expertise, extension services include crop protection, soil and fertilizer, livestock, economic management and so on. The public extension service system in China has gone through several reforms and made great progress in terms of institutional management and providing services for farmers (Gao 2008). Nevertheless, as in any public bureaucracy, because extension personnel in China are politically accountable to a large number of public servants and private commercial activities, the quality of their extension work has become a secondary priority (Hu *et al*. 2009).

As an alternative to traditional agricultural extension, the FFS approach has been promoted and expanded in many developing countries (Van den Berg and Jiggins 2007). By delivering training to a group of farmers and contracting with the most qualified one as the farmer-trainer, the FFS aims to rectify the problem of accountability and to introduce a participatory mode of extension. Farmer field schools are expected to ensure the quality and relevance of extension service provided to individual farmers.

Reforming the agricultural extension is a major part of China's recent agricultural agenda. After three years of pilot FFS projects that disseminated technology to greenhouse vegetable farmers in Beijing, the MoA has proposed the FFS as a core tool for China's agricultural extension service. Improving the efficiency of fertilizer use and pest management are major components of FFS programme.

The MoA will use the results of the effectiveness of FFS on reducing excess fertilizer use (and the associated environmental and social-economic impacts) to guide scaling up of its national FFS programme in the coming years. Since a rigorous evaluation of the FFS has not been conducted in China, we seek to fill this gap by using an RCT to evaluate the impact of FFS projects implemented by the MoA. Our findings will provide inputs to the MoA to help it decide whether and how to scale up FFS use in China.

FFS implementation locations

Our study was conducted in Anhui and Hebei provinces which are large producers of rice and tomatoes, respectively. Anhui is located in the hinterland of eastern China, with a total population of 63 million, of which 69 per cent is rural. The annual average temperature is 14–17 degrees and average annual rainfall is about 700–1,700 mm. Due to the suitable agro-climate conditions with a frost-free period of up to 200–250 days, Anhui's perennial food production ranks sixth in the country. The total crop planting area is up to eight million hectares, of which food crops account for 65 per cent. The perennial total grain production is about 25 million tonnes, of which wheat and rice account for 80 per cent. The province has more than 3,000 agricultural extension organizations with over 30,000 workers.

Hebei province is located in north China, north of the Yellow River. It is one of the most important national grain and oil production areas, with arable land up to six million hectares ranking fourth in the country. The four seasons are distinctive, and the average annual rainfall is about 350–370 mm. Hebei also has the most diversified landscape in China, including plateaus, mountains, hills, basins and plains. Due to the varying regional agro-climate conditions, many different kinds of crops grow here.

The Chinese government invests heavily in training and subsidizes fertilizer use for rice farmers, since rice is the primary grain consumed in China. In comparison, tomato is a cash crop and part of a fully liberalized market in terms of both production and marketing. Moreover, farmers can adopt a short or a long growing season for tomatoes, and we studied both types of farmers. Scientists have shown evidence of excessive N fertilizer use in both crops. By examining the differences in effects of the FFS programme on two very different crops, we can identify whether FFS works better for certain crops and make appropriate policy recommendations.

We selected counties within each province in conjunction with the MoA, based on production levels and advice from the MoA about the suitability of the counties for the study. In Anhui, the counties chosen were Tian Chang and Ju Cao, and in Hebei, we focused on Gao Cheng, Yong Qing and Rao Yang. Figure 3.1 shows a map with the two provinces chosen for our study, with the counties chosen within them.

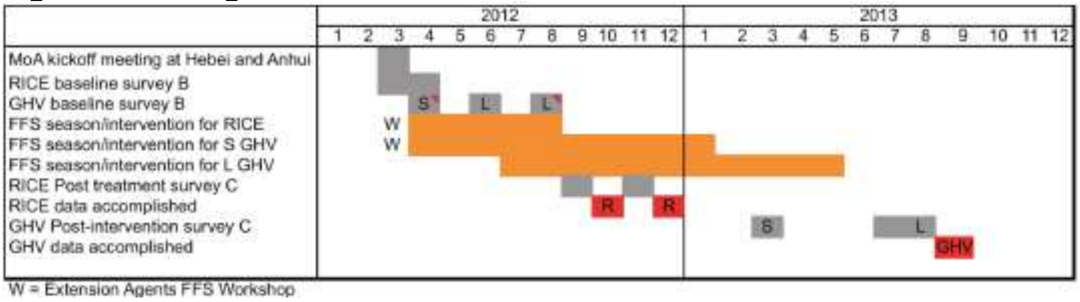
Figure 3.1: Provinces in China for FFS implementation and evaluation: Anhui and Hebei (with highlighted counties)

Source: Wikimedia Commons; includes edits by the authors

4. Linking programme implementation and impact evaluation timelines

Figure 4.1 shows the timeline for the intervention and the endline surveys. The year 2011 was spent in engaging the MoA and the local officials, designing the baseline surveys, carrying out the sample selection and conducting the baseline surveys. Figure 4.1 focuses on 2012 and 2013.

Figure 4.1. Programme timeline



W = Extension Agents FFS Workshop

Note: The area in red denotes the ending time for field work in RICE and GHV.

The extension agents were trained in March 2012. The FFS for rice farmers was implemented during the growing season of April 2012 through August 2012. Implementation for tomatoes was more complex, since tomatoes have a short (S) or a long (L) season. The FFS for the short season was implemented during the short growing season from April 2012 to the end of 2012, while the FFS for the long season for tomatoes started in mid-2012 and ended in May 2013. The endline survey for rice was completed in November 2012, while the endline surveys for short-season and long-season tomatoes were completed in March and October of 2013, respectively.

5. Methodology: evaluation design and implementation

Since the FFS programme was delivered at the village level, we designed and implemented a clustered randomized controlled trial focus on two crops. By conducting the evaluation on two very different crops, we hoped to identify crop-specificity of the effects of FFS. We did detailed power calculations for each crop separately to determine the number of villages and farmers per village. Farmers growing rice (tomatoes) during the baseline year in Anhui (Hebei) and stating intention to grow the same crop next year were deemed eligible for the study. Appendix C provides further details on the power calculation.

In Anhui (rice), we chose two counties suggested by the MoA (selected on the basis of planted area and willingness of the county to participate in the study), and four townships each from among the largest producers of rice. We randomly selected seven villages from each township for a total of 56 villages. Based on a matching algorithm that we ran based on data collected at the village level, we selected 28 villages into the treatment group (which received FFS training) and 28 into the control group (which did not receive FFS training). Our aim was to have 15 farmers randomly selected from each treatment and control village. Moreover, in treatment villages, we randomly selected 10 farmers to be ‘exposed’ farmers to study diffusion effects. This made a total of 1,120 farmers, as dictated by our power calculations.

In Hebei, we chose three counties suggested by the MoA (again chosen based on planted area and willingness of the county to participate in the study), but we were constrained by the number of villages in the selected counties. We chose 36 villages and followed our matching algorithm to assign treatment and control villages. Farmer selection in these villages mirrored Anhui (15 each in treatment and control groups and 10 in the exposed group), for a total of 720 farmers.

In practice, we chose more farmers to account for an estimated 15 per cent attrition and non-compliance. Details on recruiting farmers are given in Appendix A; we provide a brief summary here. In each treatment village, this involved recruiting 18 farmers each in the control and treatment groups and 12 in the exposed group. In the treatment villages, survey enumerators randomly selected 18 households from those eligible, and personally invited them to participate in the FFS programme by describing the nature of the programme and terms of participation.[6] Farmers had a day to decide whether or not they would like to participate, after which survey enumerators asked both refusing and accepting households to participate in a baseline survey. In subsequent rounds, additional farmers were invited equal to the number of declining farmers in the previous rounds; however, only farmers who accepted the invitation in these subsequent rounds were surveyed. This procedure allowed us to discern any systematic patterns in refusals, while minimizing the burden on enumerators and those refusing. After reaching the target of 18 households, seven more invitations were extended to fill the minimum FFS quota of 25 farmers prescribed by the MoA.

The 12 members of the exposed group in the treatment villages were also randomly selected from eligible households. Survey enumerators asked selected households whether they would be interested in participating in a survey, and willing households became part of the exposed group and were interviewed.

Control group households were surveyed using a similar process, where enumerators screened and selected households in control villages that met the eligibility criterion until they reached 18 eligible households in each village. Survey enumerators asked selected households whether they would be interested in participating in a survey, and willing households became part of the control group and were interviewed.

The balance tables for rice and tomatoes are given in Appendix E (see Tables E.1 and E.2). We compare average characteristics of households from FFS with those of non-FFS villages in terms of demographic characteristics, times of nutrient (fertilizer) and pesticide application, amount of nutrient and pesticide input, off-farm employment time, experience of rice agricultural skills training in the past three years, the number of total plots, the size of the biggest plot, cost of fertilizer and pesticides, measures of social network and so on. Equality in means between treatment and exposed groups as well as control and treatment

[6] RAND's Human Subjects Protection Committee vetted (by reviewing submitted documents and holding a hearing) the recruitment and interview protocol, as well as all the survey questionnaires.

groups cannot be rejected for almost all characteristics. In other words, our randomization seems to have worked well to produce a balanced sample.

Our analysis approach is based on the evaluation design, and we describe this approach in detail in Section 7.

6. Programme implementation

6.1 Participation in the programme

Our baseline survey was done in two stages, the first to collect demographic and other information (survey A), and the second to collect fertilizer usage (survey B, in Figure 4.1). By separating the fertilizer usage survey, we were able to account for the fertilizer used during the entire season. Farmers saved used bags of fertilizers so that the amount used could be tallied more accurately. We denote our endline survey as survey C.[7] In the implementation, the baseline survey for rice included 1,339 farmers by design, while for tomato-growing counties we included 929 farmers.

Despite best efforts on the parts of the implementing and evaluating teams, there was a sizable attrition. Tables 6.1 and 6.2 present details on attrition from our sample for rice and tomatoes, respectively. We present attrition rates across surveys as well as total attrition rates.

Table 6.1. Missing sample in survey Anhui (rice), 2011–2012

	Total	Sample by design (%)			
		T	**R**	**E**	**C**
	(N = 1,339)	**(N = 513)**	**(N = 42)**	**(N = 279)**	**(N = 505)**
Missing in baseline survey B	168	61	7	27	73
Attrition rate (%)	13	12	17	10	14
Additional missing in endline survey	148	47	5	22	73
Addl. attrition rate (%)	11	9	12	8	5
Total missing	316	108	12	49	146
Total attrition rate (%)	24	21	29	18	19

T: Treatment (accepted invitation to participate in FFS in the treatment villages)

R: Refused (did not accept invitation in the treatment villages)

E: Exposed (not randomly assigned invitation letter in the treatment villages)

C: Control (farmers in the control villages)

[7] Survey questionnaires are available from the authors upon request.

Table 6.2. Missing sample in survey Hebei (tomatoes), 2011–2013

	Total	Sample by design (%)			
		T	R	E	C
	(N = 766)	(N = 325)	(N = 1)	(N = 117)	(N = 323)
Missing in baseline survey B	79	29	0	14	36
Attrition rate (%)	10	9	0	12	11
Additional missing in endline survey	197	76	1	39	81
Addl. attrition rate (%)	26	23	100	33	25
Total missing	276	105	1	53	117
Total attrition rate (%)	36	32	100	45	36

T: Treatment (accepted invitation to participate in FFS in the treatment villages)

R: Refused (did not accept invitation in the treatment villages)

E: Exposed (not randomly assigned invitation letter in the treatment villages)

C: Control (farmers in the control villages)

As explained in Section 5, our experimental design involved issuing an invitation to participate in the FFS and a farmer could accept or refuse. Thus we distinguish between the group **T**, which accepted the invitation, and **R**, which refused. The 'exposed' group, which will be used to study diffusion effects, is selected in treatment villages from non-invited farmers and is denoted by **E**. The farmers selected in the control villages are denoted by **C**.

As seen in row 2 of Table 6.1, in the time between the two components of the baseline survey (A and B), 13 per cent of rice farmers attrited. This attrition rate does not differ significantly across the groups T, E and C. Between the baseline survey B and the endline survey in 2012, we lost an additional 11 per cent of the sample to attrition. The main explanation for this non-trivial attrition based on field inquiries appears to be the extensive amount of off-farm activities (non-agricultural jobs) in which farmers participate.

Another driver of attrition in control counties was large-scale land consolidation. In recent years, the Chinese government has promoted initiatives for land consolidation, which has resulted in consolidation in surveyed villages in the study area. Private investors rent farmland from a large number of individual farmers, some of whom had been initially selected in the control sample. Given the MoA's FFS programme, land consolidation was suspended in the treatment villages but not in the control villages. This could also influence fertilizer use in control villages independent of the FFS.

In Table 6.2, 10 per cent of tomato farmers dropped out between surveys A and B, and the attrition rate did not vary significantly across the groups T, E and C. However, between baseline survey B and endline survey C, an additional 26 per cent of the sample was lost to attrition. Based on field inquiries, the high attrition rate has three main causes. First, tomato farmers have very busy schedules, even more so than rice farmers, and especially during the growing season, so it is hard for them to guarantee attendance at the FFS. Second, the continuity of tomato farming is not as good as rice. Many surveyed plots were diverted to other crops based on the projection of farmers' market demand (despite stated intentions to continue with tomatoes, a criterion for eligibility to participate in the study).

The total attrition rate in the treatment group is only slightly higher for the treatment group for rice (Table 6.1) and higher in the control group for tomatoes (Table 6.2), which is suggestive of attrition not being directly connected to the treatment itself.

Another challenge related to how the evaluation was implemented involves additional training received by farmers in our sample that wasn't provided through the FFS. Greenhouse tomato is a cash crop, which returns a relatively high profit. The average net income of greenhouse tomato farmers can reach above 100,000 RMB per year. Farmers' ability to afford agricultural inputs such as fertilizer, pesticide and technical tools given their high incomes attracts many agricultural dealers and extension staff to hold trainings. Hence, for farmers, FFS is not the only source of agricultural information, and they could instead choose other training programmes. Some alternative training programmes, especially by agricultural dealers, could even run counter to the teachings of the FFS. Based on Table 6.3, about 40 per cent of tomato farmers in both the treatment and control groups received additional training from a non-FFS programme (either an agricultural extension agent or a commercial agricultural dealer).

Table 6.3. Non-FFS training programmes attended by tomato farmers

Attended training programmes other than FFS?		Treatment group	Control group
	No	115	129
	Yes	78	77
Source of training		78	77
	Agriculture extension	49	46
	Agricultural dealer	26	28
	Others	3	3

What are the implications of attrition? The attrition between baseline surveys A and B caused data on fertilizer usage to be incomplete. Therefore, we were unable to conduct analysis on whether fertilizer usage is systematically related to subsequent attrition or non-compliance among those who were missing in baseline B.

For farmers who dropped out from survey B to survey C, we analyzed attritors in greater detail (Tables 6.4 through 6.6). Since we still have baseline A information even for those missing in baseline B, we compared demographic information of those present in baseline B and those missing to see if there were systematic differences.

First, we compared baseline observables of attritors between the treatment and the control groups.

Table 6.4. Comparison of baseline observables of attritors between the treatment and the control groups (rice)

	Missing from treatment group	Missing from control group	Test of means
Education (years)	3.12	2.7	0.63
Age	54	52	0.53
Sex (fraction of male)	0.52	0.44	0.56
Experience of rice farming for the primary labour (years)	33	31	0.59
Fraction of participated in rice agricultural skills training in the past 3 years	0.04	0.03	0.83
N	25	34	

Table 6.4 shows that there is no systematic difference of characteristics between attritors in the treatment and control group for rice.

Table 6.5 shows that attritors in the control group for tomatoes had a higher chance of having participated in tomato agricultural skills training in the past three years. However, even this difference disappears if we do not take into account Yon Qing, a county with a problematic record of implementing the experimental design, as discussed below.

Table 6.5. Comparison of baseline observables of attritors between the treatment and the control groups (tomatoes)

	Missing from treatment group	Missing from control group	Test of means
Education (years)	7.5	8.4	0.39
Age	48	48	0.99
Sex (fraction of male)	0.8	0.9	0.33
Experience of tomato farming for the primary labour (years)	13	16	0.49
Fraction of participated in tomato agricultural skills training in the past 3 years	0.3	0.7	0.02
N fertilizer use (kg/ha)			
N	16	15	

Table 6.6. Comparison of baseline observables of attritors between the treatment and the control groups (tomatoes without Yong Qing county)

	Missing from treatment group	Missing from control group	Test of means
Education (years)	7.6	9.5	0.18
Age	49	44	0.3
Sex (fraction of male)	0.8	0.8	0.94
Experience of tomato farming for the primary labour (years)	13	12	0.85
Fraction of participated in tomato agricultural skills training in the past 3 years	0.36	0.67	0.26
N fertilizer use (kg/ha)			
N	11	6	

Next, we compare baseline observables of attritors and non-attritors as a whole. Table 6.7 shows that rice farmers who leave the sample tend to be men with lower education not inconsistent with the explanation involving off-farm activity mentioned above. The fact that attritors have less education suggests that farmers who most need the FFS training are not getting it; the impact of the programme would be higher if incentive-compatible schemes were designed to engage them in the FFS. Tables 6.8 and 6.9 show there is no systematic difference between attritors and non-attritors among tomato farmers.

Table 6.7. Comparison of baseline observables of attritors and non-attritors (rice)

	Attritors	Non-attritors	Test of means
Education (years)	2.9	4.8	0.00
Age	53	54	0.65
Sex (fraction of male)	0.48	0.75	0.00
Experience of rice farming for the primary labour (years)	32	31	0.56
Fraction of participated in rice agricultural skills training in the past 3 years	0.03	0.08	0.22
N	59	1,171	

Table 6.8. Comparison of baseline observables of attritors and non-attritors (tomatoes)

	Attritors	Non-attritors	Test of means
Education (years)	7.9	8	0.91
Age	48	45	0.17
Sex (fraction of male)	0.87	0.87	0.93
Experience of tomato farming for the primary labour (years)	14	13	0.13
Fraction of participated in tomato agricultural skills training in the past 3 years	0.52	0.44	0.4
N fertilizer use (kg/ha)			
N	31	677	

Table 6.9. Comparison of baseline observables of attritors and non-attritors (tomatoes without Yong Qing county)

	Attritors	Non-attritors	Test of means
Education (years)	8.3	8	0.66
Age	0.82	0.83	0.42
Sex (fraction of male)	47	45	0.92
Experience of tomato farming for the primary labour (years)	12	12	0.86
Fraction of participated in tomato agricultural skills training in the past 3 years	0.47	0.4	0.55
N fertilizer use (kg/ha)			
N	17	513	

As mentioned above, Tables 6.1 and 6.2 seem to suggest that treatment did not induce attrition. We now use endline attrition to test whether the probability to attrit is affected or not by the treatment.

Table 6.10 shows that the probability to attrit is actually higher for rice farmers in the control group. Tables 6.11 and 6.12 show there is no difference in the probabilities for tomato farmers.

Table 6.10. Probability of attrition that is affected by the treatment (rice)

	Treatment group	Control group	Test of means
Attrition rate	0.14	0.21	0.003
N	513	505	

Table 6.11. Probability of attrition that is affected by the treatment (tomatoes)

	Treatment group	Control group	Test of means
Attrition rate	0.29	0.3	0.76
N	325	323	

Table 6.12. Probability of attrition that is affected by the treatment (tomatoes without Yong Qing county)

	Treatment group	Control group	Test of means
Attrition rate	0.26	0.3	0.4
N	253	234	

Finally, in Tables 6.13 through 6.15 we examine attrition across the groups based on endline attrition alone, since our analysis of effect of the FFS programme will rely on the endline (and baseline) data that is actually available.

Table 6.13. Balance table for endline rice missing farmer sample

	Missing from treatment group	Treatment group	Test of means	Missing from control group	Control group	Test of means	Missing from exposed group	Exposed group	Test of means
Education (years)	3.57	4.84	0.05	4.81	4.74	0.9	4.77	5.51	0.43
Age	52.47	53.62	0.48	52.66	53.78	0.44	53.36	51.68	0.47
Sex (fraction of male)	0.45	0.24	0	0.29	0.21	0.16	0.23	0.27	0.69
Experience of rice farming for the primary labour (years)	30.09	31.31	0.52	28.89	30.87	0.25	30.27	30.05	0.93
Fraction of participated in rice agricultural skills training in the past 3 years	0.02	0.08	0.15	0.05	0.09	0.29	0.05	0.07	0.76
The rice produced from the selected plot used is for self-consumption (1 = yes)	0.6	0.44	0.04	0.66	0.45	0	0.64	0.49	0.22
N fertilizer use (kg/ha)	237.78	179.69	0	236.11	173.6	0	246.54	177.94	0
N	47	356		73	359		22	108	

Table 6.14. Balance table for tomato endline missing farmer sample

	Missing from treatment group	Treatment group	Test of means	Missing from control group	Control group	Test of means	Missing from exposed group	Exposed group	Test of means
Education (years)	8.33	8.25	0.82	8.6	7.91	0.07	7.85	8.1	0.75
Age	44.56	44.62	0.96	43.86	44.21	0.78	48.85	46.46	0.26
Sex (fraction of male)	0.66	0.68	0.74	0.8	0.78	0.7	0.74	0.85	0.27
Experience of tomato farming for the primary labour (years)	12.86	11.92	0.28	13.11	12.16	0.27	15.03	13.92	0.44
Fraction of participated in tomato agricultural skills training in the past 3 years	0.39	0.38	0.83	0.5	0.46	0.51	0.41	0.49	0.5
N fertilizer use (kg/ha)	461.39	367.97	0.04	336.64	501.3	0.01	583.71	380.65	0.1
N	76	193		81	206		39	39	

Table 6.15. Balance table for tomato endline missing farmer sample (without Yong Qing county)

	Missing from treatment group	Treatment group	Test of means	Missing from control group	Control group	Test of means	Missing from exposed group	Exposed group	Test of means
Education (years)	8.47	8.34	0.73	8.29	8.05	0.56	7.33	8.03	0.50
Age	44.53	44.51	0.99	44.05	42.89	0.41	48.67	47.30	0.62
Sex (fraction of male)	0.65	0.65	0.98	0.76	0.75	0.82	0.56	0.85	0.02
Experience of tomato farming for the primary labour (years)	12.57	11.61	0.35	13.18	11.24	0.06	14.67	14.18	0.81
Fraction of participated in tomato agricultural skills training in the past 3 years	0.37	0.35	0.74	0.47	0.45	0.79	0.39	0.45	0.66
N fertilizer use (kg/ha)	344.65	323.09	0.61	324.74	364.96	0.19	342.50	340.09	0.96
N	51	174		63	154		18	33	

Baseline nitrogen fertilizer usage by farmers who left the sample is in general significantly higher than those who stayed. In other words, farmers who most need the training are dropping out from the treatment, which could underestimate the effect of the treatment. In addition, more men seem to be missing in Anhui than in Hebei. This is consistent with the reason given above for attrition in the rice sample. Rice is not a cash crop and male farmers choose to work outside the farm. Indeed, field experience suggests that those engaged in off-farm activities tend to use excessive fertilizer at the start of the season and leave the farm. The correlation coefficient between an indicator for off-farm activity and baseline fertilizer usage is 0.32 for those missing in the sample, while it is only 0.07 for those who stayed.

The implementation of the experimental design as well as the FFS for tomatoes was most problematic (least compliant with protocols) in Yong Qing county of Hebei. In Table 6.15, we present the information in Table 6.14 after excluding the Yon Qing sample. We found that the characteristics are much more balanced across those leaving and staying in the sample across all three groups. Therefore, we deal with attrition for tomatoes (not to mention improper implementation) by also examining results by dropping the Yon Qing county sample.

6.2 Sample deviation from the experimental design

In addition to attrition, another challenge in implementation we faced was with deviation from the experimental design, which is discussed in this sub-section.

Table 6.16 shows a 'transition matrix' of how the four rice sample groups were intended to be and how they ended up being; in other words, how the intended (by design) sample breakdown differed from the eventual sample breakdown (by implementation). As mentioned above, for rice farmers, we focus only on the sample of 1,171, who were not missing in the baseline B survey. If the experiment had proceeded exactly according to design, the off-diagonal elements in the above matrix would have been zero. In the rest of this document, we concatenate the sub-group by design and sub-group by implementation to refer to the transition of a group from design to implementation. For example, R-T refers to one type of non-complying group: refused to be part of the treatment group when invited, but eventually became part of that group.

Table 6.16. Sample by design and by implementation in Anhui (rice), 2011–2012

Sample by design	N	Sample by implementation				
		T	R	E	C	Missing in endline survey
	1,171	472	51	142	359	147
T	452	356 (79%)	16 (4%)	33 (7%)	0 (0%)	47 (10%)
R	35	12 (34%)	18 (51%)	0 (0%)	0 (0%)	5 (14%)
E	252	104 (41%)	17 (7%)	108 (43%)	0 (0%)	22 (9%)
C	432	0 (0%)	0 (0%)	0 (0%)	359 (83%)	73 (17%)

T: Treatment (accepted invitation to participate in FFS in the treatment villages)
R: Refused (did not accept invitation in the treatment villages)
E: Exposed (not randomly assigned invitation letter in the treatment villages)
C: Control (farmers in the control villages)

Note: The percentages refer to the breakdown of the design groups according to how they ended up in the implementation; that is, the column percentages should add up to 100% for each row.

The Chinese MoA programme guidelines dictate that each FFS must have an enrolment of at least 25 farmers. In the initial experimental design, 79 per cent of group T participants were compliers. As such, some FFS did not have sufficient enrolment to meet programme guidelines. To comply with the MoA guidelines, additional participants were recruited randomly by sending a second round of invitation letters to farmers in group R (composed of farmers who refused to participate in FFS when the first-round invitations were sent) and group E (composed of farmers who did not receive an invitation to participate in FFS in the first round). Individuals who accepted the second-round invitations from group R and group E are denoted R-T and E-T, respectively. Twenty-eight per cent of group R individuals and 37 per cent of group E individuals agreed to participate in FFS after the second-round invitations were sent.

Farmers who converted themselves from group T to group R or group E are denoted T-R or T-E. In the implementation, 4 per cent of group T farmers, those who accepted the early invitations, somehow refused to attend FFS when extension staff reached out to them. Seven per cent of group T farmers could not be reached by extension staff, but they were surveyed and categorized in group E. Because the T-E farmers received and accepted early invitation while farmers in initial group E didn’t, the T-E farmers could be a potential source of bias in measuring programme impact.

Table 6.17 shows how the intended (by design) tomato sample breakdown differed from the eventual sample breakdown (by implementation). As with the rice FFS, to comply with the enrolment requirement of the MoA, additional participants for rice were recruited. But unlike rice, in addition to the farmers converted from group E, extension staff also recruited farmers who were not even considered for the programme initially. Since these farmers did not take the baseline survey, we do not include them in the analysis.

The attrition rate varies by group. The attrition rates are very close for groups T and C, suggesting random attrition. The attrition rates are higher in groups R and E (see discussion above); however, it must be noted that their sample sizes are also much smaller.

Table 6.17. Household sample of RCT by design and by implementation in Hebei (tomato), 2011–2013.

Sample by design	N	Sample by implementation				
		T	R	E	C	Missing in endline survey
	687	219	4	64	206	147
T	296	193 (65%)	2 (1%)	25 (8%)	0	73 (25%)
R	1	0	0	0	0	1 (100%)
E	103	23 (22%)	2 (2%)	39 (38%)	0	39 (38%)
C	287	0	0	0	206 (72%)	81 (28%)

T: Treatment (accepted invitation to participate in FFS in the treatment villages)
R: Refused (did not accept invitation in the treatment villages)
E: Exposed (not randomly assigned invitation letter in the treatment villages)
C: Control (farmers in the control villages)

Note: The percentages refer to the breakdown of the design groups according to how they ended up in the implementation; that is, the column percentages should add up to 100% in each row.

The remaining analysis uses the sample of 973 (T + C + E by implementation) rice farmers and the sample of 489 (T + C + E by implementation) tomato farmers for whom baseline B and endline C data exists (or a subset of these in some cases). As discussed below, we deal with the deviation from design by comparing means between treatment and control groups for a range of cases: the most compliant groups (treatment-on-treated analysis) to the most inclusive groups, as randomized (intent-to-treat analysis). In our regressions, where we use the whole sample, we instrument for treatment/participation using the invitation to participate.

7. Impact analysis and results of the key evaluation questions

The overall objective is to study the effect of FFS on sub-optimal (*a priori*, excessive) fertilizer use. In this section, we focus on this primary aim and discuss diffusion in the next section. We expect significant contamination effects on the exposed group because of the unintentional deviation between the experimental design and implementation discussed in detail in the previous section.

Given the non-trivial crossover across the groups brought about by the FFS implementation (as discussed in Section 6.2), we start with a comparison of T-T and C-C, which offers the simplest and cleanest comparison. This will constitute an analysis of the treatment on the treated. We then compare the most inclusive treatment group (which includes farmers transiting from other groups into treatment) and the control group. In the Appendix we present results for groups with an intermediate level of compliance/inclusion into the treatment group.

Our analysis plan follows from our RCT evaluation design, starting with comparing mean outcomes for the treatment and control groups. We also draw on supplemental methods to address differential responses between low and high fertilizer users and non-compliance in treatment assignment. Since the participation or non-participation of selected farmers in the FFS does not appear to have been driven purely by characteristics inherent to the programme as much as the way the design was implemented in the field, we rely primarily on the above-mentioned methods of analysis to understand programme effectiveness. However, we also present an intent-to-treat analysis at the end, comparing the treatment and control groups as randomized. In summary, our analysis approach includes the following, which we explain in greater detail throughout Section 7:

- Comparison of means for pure treatment and control groups
- Comparison of means for fertilizer use sub-groups
- Difference-in-difference with and without controls to assess distance from the agronomic optimum
- Instrumental variable (IV) estimates to assess non-compliance
- Intent-to-treat analysis

7.1 Treatment group (T-T) vs control group (C-C) results

We first evaluate the effect of FFS on fertilizer use by comparing the change in the mean fertilizer use between the pure (that is, complying) treatment and pure control groups.

7.1.1. Difference in means

We compare the differences in mean fertilizer use between groups T-T (those treated by implementation) and group C-C (control groups) for rice in Table 7.1.

Table 7.1. Effect of FFS on chemical fertilizer use (rice)

	Sample	Nitrogen			Potassium		
		Baseline	Endline	Delta	Baseline	Endline	Delta
Treatment (T-T)	356	180	147	−32	32	46	14
Control (C-C)	359	174	137	−37	35	43	8
Difference b/w T-T and C-C		6	10	5	−3	3	6*
p-value		0.33	0.06	0.45	0.2	0.3	0.05

*Note: * p < 0.05, ** p < 0.01*

Endline nitrogen fertilizer use reduced dramatically in *both* treatment group and control group. One main reason is likely to be the unexpected weather pattern in Anhui in the baseline survey year 2011. From June to October of 2011, rainfall was much higher than usual. Cloudy and wet weather reduces photosynthesis and further leads to insufficient tiller development, which affects the growth of rice.[8] Meanwhile, long-term immersion in the water made land heating impossible, so that the root of rice rotted in such a damp and oxygen-deficient environment leading to a crop disease called lodging. In addition, excessive nitrogen fertilizer application aggravates lodging since unbalanced nutrients reduce the capacity of rice to survive when exposed to extreme weather. Because farmers applied fertilizer in June and July and did not realize the impact of excessive rainfall until harvest, their usage of fertilizer in baseline was not influenced by unusual weather. However, it is common knowledge among farmers that lodging can be addressed by reducing fertilizer use, so in the endline year 2012, farmers reduced fertilizer use in general due to concerns about potential rainy weather. This common trend across both treatment and control groups explains why fertilizer reduction happened simultaneously in both groups.

Although there is no significant difference in the change in average nitrogen application between T-T and C-C ($p = 0.45$), as we discuss below, there are heterogeneous effects, which shows that the impact of the FFS varied across fertilizer use sub-groups.

The application of K fertilizer increased post-intervention in both treatment and control groups, with the treatment group experiencing a larger increase in K fertilizer use ($p = 0.05$). It is also common knowledge among farmers that K fertilizer can increase the resistance of rice crops against lodging, even if only by a small amount. However, most farmers focus on nitrogen fertilizer application because K fertilizer is expensive—so much so that some farmers do not use it at all.

[8] The tiller is the stage of the rice plant that follows the seedling.

It is possible that farmers increased K fertilizer use to prevent losses caused by lodging and the possibility of unusual weather in the endline year. This may explain why farmers in both treatment and control groups increased K fertilizer use. On the other hand, given the relatively low level of local K fertilizer use, teaching farmers about the benefits of using more K fertilizer was one of the main goals of the FFS curriculum. The differential change in mean K fertilizer use between the treatment and control groups could be attributable to the FFS intervention.

Table 7.2. Effect of FFS on chemical fertilizer use in tomato planting

	N	Nitrogen			Potassium		
		Baseline	Endline	Delta	Baseline	Endline	Delta
Treatment (T-T)	193	368	482	114	456	591	134
Control (C-C)	206	501	488	−13	628	588	−40
Difference b/w T-T and C-C		−133*	−6	127**	−171*	3	174**
p-value		0.03	0.86	0.01	0.03	0.96	0.01

Note: * *p < 0.05*, ** *p < 0.01*

Table 7.2 presents the analogue of Table 7.1 for tomatoes. In the endline year, N fertilizer use greatly increased in the treatment group while it actually slightly decreased in the control group. Potassium fertilizer usage shows similar trends. Since the baseline fertilizer usage in the treatment group was significantly lower, both groups ended up with very similar usage in the endline.

Since fertilizer usage is highly heterogeneous among farmers, a direct comparison of averages between the T-T and C-C group might mask one of the main expected contributions of the FFS, which is to educate the farmers about optimal fertilizer usage. The average suggested optimum for N fertilizer by agricultural experts for growing rice is 165–180 kg/ha under normal weather conditions. In other words, one of the anticipated effects of the FFS is for farmers whose fertilizer use is below the optimal level to increase usage while those who are using excessive fertilizer should reduce the application. If this indeed happened, the distribution of fertilizer use in the treatment group should be closer to the optimal level in the endline year than in the baseline year when compared with the corresponding difference in the control group.

To further explore the heterogeneity in fertilizer usage, we break N fertilizer usage by quintile. This is shown for rice in Table 7.3.

Table 7.3. Comparison of N fertilizer usage by quintile (rice)

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
TT	356			n = 72	n = 71	n = 71	n = 71	n = 71
		Nitrogen baseline[1]	180	69	124*	164	223*	320**
		Nitrogen endline[2]	148	114	124	142	152	206
		Delta of nitrogen[3]	–32	45*	0	–22	–71	–114
		% change	–18%	65%	0	–13%	–32%	–36%
CC	359			n = 72	n = 72	n = 72	n = 72	n = 71
		Nitrogen baseline	174	75	128	167	213	287
		Nitrogen endline	137	95	114	132	154	190
		Delta of nitrogen	–37	20	–14	–35	–59	–97
		% change	–21%	27%	–11%	–21%	–28%	–34%

Note: * *$p < 0.05$*, ** *$p < 0.01$*
[1] t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
[2] t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
[3] t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

In Table 7.3, nitrogen use increased in the first quintile (0–20 per cent) in both control and treatment groups while nitrogen use reduced for the other quintiles (except quintile 20–40 per cent in treatment group, where it stayed the same). The reduction is the highest in the top quintile compared to other quintiles in both treatment and control groups. However, the increase in the first quintile for the treatment group (45) is substantially higher than in the control group (20, p = 0.03), while the decreases in the top two quintiles are slightly higher and not significant.

As mentioned earlier, increasing potassium usage was also one of the aims of the rice FFS. Table 7.4 shows potassium usage, also by quintile.

Table 7.4. Comparison of K fertilizer usage by quintile (rice)

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
TT	356			n = 114	n = 29	n = 76	n = 76	n = 61
		Potassium baseline[1]	32	0	13*	28**	46*	87
		Potassium endline[2]	46	33	38	46	50	68
		Delta[3]	14	33	25	18	4*	−19
		% change	44%	–	192%	64%	9%	−22%
CC	359			n = 95	n = 52	n = 74	n = 75	n = 63
		Potassium baseline	35	0	17	31	48	93
		Potassium endline	43	34	36	47	43	54
		delta	8	34	19	16	−5	−39
		% change	23%	–	112%	52%	−10%	−42%

Note: * $p < 0.05$, ** $p < 0.01$

[1] t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.

[2] t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.

[3] t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

Potassium use increased dramatically in the first quintile (0–20 per cent) in both control and treatment groups. The use of potassium reduced significantly in both groups in the top quintile 80–100 per cent, but reduction in the control group is much higher as compared to the treatment group. It is useful to reiterate the priors of agronomists that K use is too low for most farmers, and one of the aims of the FFS was to increase it on average.

We also break down fertilizer usage by quintile for tomatoes. This is shown in Tables 7.5 and 7.6. To build the foundation of greenhouses, farmers have to remove the surface of the soil, which contains more nutrients and is more suitable for farming. Therefore, a great amount of fertilizer has to be applied to the deep-level soil to make it arable. This explains why the total amount of fertilizer in tomato is higher than that in rice.

Table 7.5. Comparison of N fertilizer usage by quintile (tomato)

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
				n=39	n=39	n=38	n=39	n=38
TT	193	nitrogen baseline[1]	368	108**	218**	291**	428**	803*
		nitrogen endline[2]	482	441	332	463	450	729
		delta of nitrogen[3]	114	333	114	172*	22	-74*
		% change	31%	308%	52%	59%	5%	-9%
				n=42	n=41	n=41	n=41	n=41
CC	206	nitrogen baseline	501	155	269	367	492	1233
		nitrogen endline	489	399	413	444	470	719
		delta of nitrogen	-13	244	144	77	-22	-514
		% change	-3%	157%	54%	21%	-4%	-42%

Note: * $p < 0.05$, ** $p < 0.01$
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

In Table 7.5, nitrogen use increased in the first quintile (0–20 per cent) in both control and treatment groups, but the increase in the treatment group is significantly higher than that in the control group. The reduction is the highest in the top quintile compared to others in both treatment and control groups. Since the mean of the top quintile in the control group is much higher than that in the treatment group in the baseline year, although both groups reduced to around 720 kg/ha in the endline year, the reduction in the top quintile of the control group is substantially higher than that in the treatment group. Similar trends can be found in Table 7.6.

Table 7.6. Comparison of K fertilizer usage by quintile (tomato)

Group	N		Mean	Quintile				
				0–20	20-40	40–60	60–80	80–100
TT	193			n=39	n=39	n=38	n=39	n=38
		potassium baseline[1]	456	123**	252**	361**	499**	1058*
		potassium endline[2]	591	398	509	571	578	904
		delta of potassium[3]	135	275	257	210*	79	-154*
		% change	30%	224%	102%	58%	16%	-15%
CC	206			n=42	n=41	n=41	n=41	n=41
		potassium baseline	628	170	315	436	602	1627
		potassium endline	588	524	464	431	701	821
		delta of potassium	-40	354	149	-5	99	-806
		% change	-6%	208%	47%	-1%	16%	-50%

Note: * $p < 0.05$, ** $p < 0.01$
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

As mentioned in Section 6, high-income tomato farmers attract many agricultural dealers and extension staff to hold trainings. Hence, for farmers, FFS is not the only source of agricultural information; indeed some of the training, especially by agricultural dealers, could even run counter to the teachings of the FFS. This could be a potential source of confounding reflected in the tomato results, especially at the upper end of fertilizer usage.

As mentioned in Section 6, the implementation of the experimental design as well as the FFS for tomatoes was most problematic (least compliant with protocols) in Yong Qing county of Hebei. In Table 7.7, we present the information in Table 7.5 (which includes Yon Qing) for N fertilizer usage after excluding the Yon Qing sample. We see that the reduction of N fertilizer in the treatment group is larger for the highest quintile than in Table 7.5, but there is still no effect of the FFS treatment for this quintile. The effect of FFS in increasing the fertilizer usage in the lowest quintile remains.

Table 7.7. Comparison of N fertilizer usage by quintile (tomato)(without Yong Qing county)

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
				n=35	n=35	n=35	n=35	n=34
		nitrogen baseline[1]	323	101**	207**	270**	382**	665
TT	174	nitrogen endline[2]	431	450	340	421	410	536
		delta of nitrogen[3]	108	349	133	151	28	-129
		% change	33%	346%	64%	56%	7%	-19%
				n=31	n=31	n=31	n=31	n=30
		nitrogen baseline	365	145	246	326	432	686
CC	154	nitrogen endline	415	422	359	394	424	478
		delta of nitrogen	50	277	113	68	-8	-208
		% change	14%	191%	46%	21%	-2%	-30%

Note: * $p < 0.05$, ** $p < 0.01$
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

As we mentioned earlier, farmers can adopt a short or a long growing season for tomatoes. Short-season growers rely on conventional greenhouses, which help regulate temperature but cannot be used consistently for year-round growing. Long-season farmers use a modified greenhouse referred to as a shack, which is partially below ground. This increases the temperature in the greenhouse and lengthens the growing season, which affects the recommended nutrient application rates. We divided the farmers by growing seasons and broke down the fertilizer use by quintile to see how the results change. Since the sample size of the long growing season is too small, we only looked at the short growing season here. Based on the results presented in Table 7.8, we can see that the highest quintile sees a decrease that is about the same in the treatment group as in the control group (but no higher). In the lowest quintile, the fertilizer increase in the treatment group continues to be much larger than that in the control group.

Table 7.8. Comparison of N fertilizer usage by quintile in short growing season tomato planting (without Yong Qing county)

Group	N	Nitrogen	Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
				n=31	n=30	n=31	n=30	n=30
		Baseline[1]	321	103**	211**	270**	378**	652.367
TT	152	Endline[2]	404	423	296	434	403	459.55
		Delta[3]	83	320	85	164	25	-192.82
		% change	26%	311%	40%	61%	7%	-30%
				n=29	n=28	n=28	n=28	n=28
		Baseline	369	147	253	332	436	688
CC	141	Endline	413	429	356	386	418	474
		Delta	43	282	104	53	-18	-212
		% change	12%	192%	41%	16%	-4%	-31%

Note: * $p < 0.05$, ** $p < 0.01$
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

7.1.2. Difference-in-Difference in distance from the optimum

The results of the above tables suggest that the effect of the FFS might have been to increase the usage of fertilizer among the lowest quintiles and decrease it at the highest quintile (for rice). This leads us to examine the change in the distance from the optimum fertilizer usage, which accounts for differential responses by farmers along the fertilizer use distribution.

Distance here is defined as the absolute distance from optimum range of the 165–180 kg/ha for rice and 300–375 kg/ha for the short-growing season tomato determined by agronomists for N fertilizer usage. We examine whether this distance reduced between the endline and baseline more for the treatment than the control group.

While the treatment and control groups are well balanced, one of the main contributions of the study is documenting how much heterogeneity exists in fertilizer usage, which was uncovered only after the baseline survey. Therefore the distribution of farmers by baseline fertilizer usage is not fully balanced. Using differences in our regressions and quintile analysis in comparisons of means (done earlier) addresses this issue.

The first column of Table 7.9 shows the regression of the differences in distance from the optimum between endline and baseline on participation for rice farmers.

Participation in FFS is significantly and negatively associated with this difference in distance.

Table 7.9. Regression of differences in distance from optimum (rice)

	(1)	(2)
FFS Treatment	-15.731***	-14.100**
	(-3.33)	(-2.77)
Education		0.451
		(0.65)
Female		1.139
		(0.18)
Years farming rice		0.127
		(0.57)
Organic		7.096
		(0.64)
Own consumption		-8.841
		(-1.59)
Time on off-farm work		0.037
		(0.37)
Number of total plots		1.165
		(1.52)
Mobile phone use		10.146
		(1.75)
Work with other farmers		8.024
		(0.93)
Cost of fertilizer		-6.439*
		(-2.19)
Cost of pesticide		0.019
		(1.32)
_cons	7.926**	-2.381
	(2.38)	(-0.17)
F-test	11.07	2.72
N of obs	715	656

Note: *Organic* is an indicator variable that is 1 if the farmer uses organic fertilizer; *own consumption* is an indicator variable if the farmer's household consumes the crop grown on the plot; * p < 0.05, ** p < 0.01

As discussed in Section 5, the treatment and control groups were not balanced across a few dimensions. In the second column we control for these and a few other variables. While the magnitude of the coefficient decreases slightly, it remains negative and significant.[9]

Table 7.10. Regression of differences in distance from optimum (tomato) without Yong Qing county

	(1)	(2)
FFS participation	-0.797	5.557
	(-0.02)	(0.17)
Age		-2.986
		(-1.51)
Education		-1.8
		(-0.28)
Male		27.467
		(0.73)
Years farming vegetable		3.237
		(1.13)
Time on off-farm work		0.473
		(0.85)
Other training		63.893
		(1.75)
Constant	28.655	83.523
	(1.2)	(0.82)
F-tests	0	0.87
N	293	293

Note: * p < 0.05, ** p < 0.01

For tomato farmers, the regression in the first column of Table 7.10 shows that participation in FFS is negatively associated with the reduction of distance from the optimum, but the relationship is not significant. The second regression shows that participation in FFS is associated with an *increase* of the distance from the optimum after controlling for the few variables not balanced between treatment and control groups and other variables; however, this is not significant.

[9] Only the cost of pesticide is unbalanced between the treatment and control groups. We also include the cost of fertilizers given its relevance in a regression on fertilizer usage. It enters negatively and significantly, as one might expect. Dropping this regressor does not appreciably alter the results (this is true for regressions presented later as well). Also, when we control for the distance from the optimum in the baseline (to allow for potential lack of balance between the groups and the unusual weather in the baseline), participation is still significantly negative at the 10 per cent level, showing that the FFS was effective in reducing the distance from optimal fertilizer usage.

Taken together, these regressions seem to confirm the results from the quintile analysis that the FFS intervention was more effective in reducing the distance to the optimum prescribed fertilizer amount for rice than for tomatoes.

7.2. Treatment with non-compliers analysis (T-T + E-T + R-T)

In this section we expand the treatment group of rice farmers to include non-compliers: farmers who were in the exposed group but participated in the treatment (E-T) and farmers who refused to participate initially but later took part in the FFS (R-T).[10] Hence, this is the most inclusive treatment group we consider.[11] Looking at non-compliers allows us to expand the sample and check whether the initial results are robust to inclusion of additional groups of FFS farmers. We analyze summary data from the experiment and then use IV methods in an encouragement design framework to address self-selection.

7.2.1. Difference in means

In Table 7.11, we compare the differences in mean fertilizer use between groups T-T + E-T + R-T (those treated by implementation) and group C-C (the controls) for rice and tomato farmers. In this scenario, we see results similar to those in the previous two scenarios, so we further break down the fertilizer use by quintile to examine the heterogeneity of fertilizer use. Results in the quintile comparison are similar as well. Nitrogen use for rice farmers increased in the first quintile (0–20 per cent) in both control and treatment groups while nitrogen use reduced for the other quintiles (Table 7.12). The reduction is the highest in the top quintile as compared to others in both treatment and control groups. However, the increase in the first quintile and decrease in the top two quintiles of treatment group are higher than those in the control group. Potassium use increased dramatically in the first quintile (0–20 per cent) and reduced a lot in the top quintile (80–100 per cent) in both control and treatment groups (Table 7.13). Overall, it again implies that farmers in the treatment group use more K fertilizer than those in the control group.

Our findings in the previous section for rice—the effect of the FFS is to significantly increase fertilizer usage at the lowest quintile of the distribution and decrease it slightly at the highest quintile—appear to be robust to allowing for potential contamination and choosing the most inclusive treatment group.

[10] The R-T group declined the invitation (when the programme was introduced and explained to them before it was implemented) but eventually recruited as FFS graduate. It is not entirely clear why they ended up in the treatment group, but we suspect they participated out of peer pressure.

[11] Appendix G provides additional results for an intermediate level of inclusiveness that does not include the R-T group.

Table 7.11. Comparison of means for rice: (TT+ET+RT) v CC

	Treatment (T-T + E-T + R-T)	Control (C-C)	Difference	p-value
Nitrogen baseline	186	174	12*	0.04
Nitrogen endline	148	137	11*	0.04
Delta in nitrogen	-38	-37	-1	0.83
Potassium baseline	33	35	-2	0.4
Potassium endline	47	43	4	0.1
Delta in potassium	14	8	6*	0.04

*Note: * p < 0.05, ** p < 0.01*

Table 7.12. Comparison of means for tomato: (TT+ET+RT) v CC

	Treatment (T-T + E-T + R-T)	Control (C-C)	Difference	p-value
Nitrogen baseline	369	501	-133**	0.00
Nitrogen endline	500	489	11	0.77
Delta in nitrogen	131	-13	144**	0.00
Potassium baseline	467	628	-161	0.00
Potassium endline	603	588	15	0.76
Delta in potassium	136	-40	176*	0.01

*Note: * p < 0.05, ** p < 0.01*

Table 7.13. Comparison of N fertilizer usage by quintile for rice

Group	N		Mean	Quintile 0–20	20–40	40–60	60–80	80–100
				n=95	n=94	n=96	n=93	n=94
		nitrogen baseline[1]	185	73	130	173*	234**	323**
TT	472	nitrogen endline[2]	149	108	126	143	157	208
		delta of nitrogen[3]	-36	35	-4	-30	-77	-115
		% change[4]	-19%	48%	-3%	-17%	-33%	-36%
				n=72	n=72	n=72	n=72	n=71
		nitrogen baseline	174	75	128	167	213	287
CC	359	nitrogen endline	137	95	114	132	154	190
		delta of nitrogen	-37	20	-14	-35	-59	-97
		% change	-21%	27%	-11%	-21%	-28%	-34%

Note: * *p < 0.05*, ** *p < 0.01*
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

Table 7.14. Comparison of N fertilizer usage by quintile for tomato

Group	N		Mean	Quintile 0–20	20–40	40–60	60–80	80–100
				n=44	n=43	n=43	n=43	n=43
		Nitrogen baseline[1]	369	110	217	294*	428*	801*
TT	216	Nitrogen endline[2]	500	443	346	454	487	772
		Delta[3]	131	333	129	159*	59	-29*
		% change	36%	303%	60%	54%	14%	-4%
				n=42	n=41	n=41	n=41	n=41
		potassium baseline	501	154	269	367	491	1233
CC	206	potassium endline	489	399	413	444	470	719
		delta	-13	244	144	77	-21	-514
		% change	-3%	158%	53%	21%	-4%	-42%

Note: * *p<0.05*, ** *p<0.01*
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.

3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

Table 7.15. Comparison of N fertilizer usage by quintile for tomato (without Yong Qing county)

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
				n=38	n=38	n=38	n=38	n=37
		Nitrogen baseline[1]	324	103	208	274*	384*	660
TT	189	Nitrogen endline[2]	441	443	355	440	426	545
		Delta[3]	117	339	147	166*	42	-116
		% change	36%	328%	70%	61%	11%	-18%
				n=31	n=31	n=31	n=31	n=30
		Nitrogen baseline[1]	365	145	246	326	432	686
CC	154	Nitrogen endline[2]	415	422	359	394	424	478
		Delta[3]	50	277	113	68	-9	-208
		% change	14%	191%	46%	21%	-2%	-30%

Note: * $p < 0.05$, ** $p < 0.01$
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

Table 7.16. Comparison of K fertilizer by quintile for rice

Group	N		Mean	Quintile				
				0−20	20−40	40−60	60−80	80−100
				n=140	n=49	n=107	n=94	n=82
		potassium baseline[1]	33	0	15	30	48	88
TT	460	potassium endline[2]	47	34	38	44	55*	70
		delta[3]	14	34	23	14	7*	-18
		% change[4]	44%	-	153%	47%	15%*	-20%
				n=95	n=52	n=74	n=75	n=63
		potassium baseline	35	0	17	31	48	93
CC	359	potassium endline	43	34	36	47	43	54
		delta	8	34	19	16	-5	-39
		% change	23%	-	112%	52%	-10%	-42%

Note: * *p < 0.05*, ** *p < 0.01*
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

7.2.2. IV regressions

Given the imperfect treatment group and the presence of non-compliers, we use instrumental variable regressions to address possible selection bias. We use an indicator variable, which indicates whether a participant was invited to participate in the treatment, as an IV for FFS attendance. DID in distance from the optimum, as in Table 7.9, is used as the dependent variable. While this IV might address selection from the E group, it might not address selection from the R group (though the R-T group is rather small).

In the first-stage regression, the variable of invitation is significantly associated with actual participation in FFS. In the second stage (Tables 7.17 and 7.18), after controlling for unbalanced variables and household characteristics, participation in FFS is significantly and negatively to a distance from the optimum fertilizer use. As before, for rice, participation in FFS reduces the distance from the optimum relative to the control group. Indeed, the coefficient and significance improved relative to those in Table 7.9.

Table 7.17. IV (2SLS) regression of DID in distance from optimum (rice)

DID Distance	IV
FFS Treatment	-16.770***
	(-3.32)
Education	0.395
	(0.61)
Female	-1.222
	(-0.20)
Years farming rice	0.185
	(0.9)
Organic	2.449
	(0.24)
Own consumption	-7.997
	(-1.53)
Time on off-farm work	-0.021
	(-0.23)
Number of total plots	1.025
	(1.42)
Mobile phone use	12.059*
	(2.26)
Work with other farmers	7.519
	(0.96)
Cost of pesticide	0.02
	(1.46)
Cost of fertilizer	-5.843*
	(-2.05)
_cons	0.047
	0
F-test	3.27
Number of obs	758

Note: * *p<0.05, ** p<0.01*

Table 7.18. IV regression of DID in distance from optimum for tomato (without Yong Qing county)

Instrumental variables (2SLS) regression	
DID distance	IV
FFS participation	24.035
	(0.68)
Age	-2.483
	(-1.28)
Education	-1.33
	(-0.21)
Male	32.724
	(0.89)
Years farming vegetable	3.593
	(1.27)
Time on off-farm work	0.409
	(0.74)
Other training	55.619
	(1.55)
Constant	50.616
	(0.5)
F-tests	0.8
N	306

Note: * *p<0.05, ** p<0.01*

7.3. Impact of FFS on yield

Whether FFS graduates have higher net yields and incomes (or at the least not lower amounts) and are better able to resolve farming production problems in the context of an altered fertilizer regimen, are a couple of concerns we turn to next. In the three scenarios for rice (Table 7.19 through Table 7.21), yield in the endline year increased greatly and by almost equivalent amounts in both treatment and control groups. One possible reason for growth in both groups is that the unusual weather in the baseline year influenced the yield, so with the return of normal weather, the yield jumped back to the normal level as well. In the two different scenarios for tomatoes (Table 7.22 and Table 7.23), even though the yield of the treatment and control groups both increase in the endline survey, the growth in yield of the treatment group is much higher than the control group.

Table 7.19. Rice yield comparison: T-T v C-C

	Treatment (T-T)	Control (C-C)	Difference	p-value
Yield baseline	7581	7456	124	0.23
Yield endline	8252	8134	118	0.24
Delta in yield	672	678	-6	0.95

Note: * *p<0.05*, ** *p<0.01*

Table 7.20. Rice yield comparison: (T-T+E-T) v C-C

	Treatment (T-T + E-T)	Control (C-C)	Difference	p-value
Yield baseline	7630	7456	173	0.07
Yield endline	8300	8134	166	0.07
Delta in yield	670	678	-8	0.93

Note: * *p<0.05*, ** *p<0.01*

Table 7.21. Rice yield (kg/ha) comparison: (T-T+E-T+R-T) v C-C

	Treatment (T-T + E-T + R-T)	Control (C-C)	Difference	p-value
Yield baseline	7630	7456	174	0.07
Yield endline	8301	8134	167	0.07
Delta in yield	670	678	-7	0.94

Note: * *p<0.05*, ** *p<0.01*

Table 7.22. Tomato yield (kg/ha) comparison: T-T v C-C

	Treatment (T-T)	Control (C-C)	Difference	p-value
Yield baseline	70950	79902	-8952	0.08
Yield endline	84610	83916	694	0.87
Delta in yield	13290	4321	8969	0.09

Note: * *p<0.05, ** p<0.01*

Table 7.23. Tomato yield (kg/ha) comparison: (T-T+E-T) v C-C

	Treatment (T-T + E-T)	Control (C-C)	Difference	p-value
Yield baseline	71904	79902	-7999	0.11
Yield endline	86621	83916	2704	0.50
Delta in yield	14383	4321	10062	0.05

Note: * *p<0.05, ** p<0.01*

We next explore the heterogeneous effect of FFS on yield in greater detail, particularly whether for the highest quintile the reductions in fertilizer usage have any negative effect on yield. Based on the figures in Table 7.24,which presents rice yield (in kilograms per hectare) by quintile of fertilizer use, yield increases in all of the quintile bins, especially in the first two quintile bins (0–20% and 20–40 per cent) presumably due to the increased usage of fertilizer. The highest quintile bin had the greatest reduction of fertilizer use, but yield was not negatively impacted by this decrease.

Table 7.24. Comparison of rice yield (kg/ha) by quintile fertilizer use

Group			Mean	Quintile 0–20	20–40	40–60	60–80	80–100
TT			n=356	n=72	n=71	n=71	n=71	n=71
	N fertilizer	nitrogen baseline	180	69	124	164	223	320
		nitrogen endline	148	114	124	142	152	206
		delta of nitrogen	-32	45	0	-22	-71	-114
		% change	-18%	65%	0	-13%	-32%	-36%
	Yield	yield baseline	7581	7397	7330	7210	7875	8095
		yield endline	8252	8416	8028	8055	8576	8184
		delta of yield	671	1019	698	845	701	89
		% change	9%	14%	10%	12%	9%	1%
CC			n=359	n=72	n=72	n=72	n=72	n=71
	N fertilizer	nitrogen baseline	174	75	128	167	213	287
		nitrogen endline	137	95	114	132	154	190
		delta of nitrogen	-37	20	-14	-35	-59	-97
		% change	-21%	27%	-11%	-21%	-28%	-34%
	Yield	yield baseline	7456	7167	7368	7792	7542	7413
		yield endline	8134	8220	8147	8198	8052	8052
		delta of yield	678	1053	779	406	510	639
		% change	9%	15%	11%	5%	7%	9%

Based on the figures in Table 7.25, which presents tomato yield (in kilograms per hectare) by quintile of fertilizer use, treatment group yield increases in all of the quintile bins, and the increase in the treatment group is significantly higher than that in the control group for each bin except the quintile 20–40 per cent. Yield in the top quintile of the treatment group shows a 12 per cent increase despite the greatest reduction of fertilizer use, while yield in the control group decreased slightly (-1 per cent) at the same time that fertilizer use fell dramatically (–42 per cent).

Table 7.25. Comparison of tomato yield (kg/ha) by quintile fertilizer use

Group			Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
TT				n=39	n=39	n=38	n=39	n=38
	N fertilizer	nitrogen baseline	368	108	218	291	428	803
		nitrogen endline	482	441	332	463	450	729
		delta of nitrogen	114	333	114	172	22	-74
		% change	31%	308%	52%	59%	5%	-9%
	Yield	yield baseline	70950	61958	63545	73147	72106	84398
		yield endline	84610	73172	79285	84069	92290	94739
		delta of yield	13660	11214	15740	10922	20184	10341
		% change	19%	18%	25%	15%	28%	12%
CC				n=42	n=41	n=41	n=41	n=41
	N fertilizer	nitrogen baseline	501	155	269	367	492	1233
		nitrogen endline	489	399	413	444	470	719
		delta of nitrogen	-12	244	144	77	-22	-514
		% change	-2%	157%	54%	21%	-4%	-42%
	Yield	yield baseline	79902	88449	73228	66997	72320	98310
		yield endline	83916	84930	91524	70126	74962	97678
		delta of yield	4014	-3519	18296	3129	2642	-632
		% change	5%	-4%	25%	5%	4%	-1%

7.4. Impact of FFS on knowledge score

The FFS training focuses not only on reducing excessive fertilizer use, but also promoting environmentally sound practice in general, like crop protection, scientific cultivation, and enhancing the environmental and ecological awareness of farmers. The effectiveness of the curriculum in bringing these benefits can be tested using the questions we included in the surveys. We conducted a detailed comparison of knowledge scores between pre- and post-intervention surveys by group.

Based on the statistical tests (see Table 7.26a/b), farmers in the treatment group (no matter whether the farmer is a complier or not) get a significantly higher knowledge score than those in the control group. More specifically, fertilizer, pest, cultivation and environment sub-scores for the treatment group are all higher than those of the control group. In addition, we see no significant difference between knowledge scores of farmers in the exposed group and control group.

Table 7.26a. Difference of farmer test score between 2011 and 2012 in Anhui, China

		Sample	Total score[5]	Fertilizer[6]	Pest[6]	Cultivation[6]	Environment[6]
1	Total	1171	142	37	34	53	19
2	**FFS graduates in treatment villages**[1]	**472**	**143**	**38**	**33***	**53**	**19**
3	T-T	356	143	38	33	53	19
4	R-T[2]	12	131	39	29	45	19
5	E-T[2]	104	143	38	33	53	19
6	**Non-compliance farmers in treatment villages**[1]	**51**	**123****	**32**	**30****	**48***	**13**
7	T-R[3]	16	130	34	32	49	14
8	R-R	18	133	32	32	51	18
9	E-R[3]	17	124	34	29	46	15
10	**Exposed farmers in treatment villages**[1]	**141**	**138**	**36**	**32*****	**52**	**18**
11	T-E[4]	33	130	34	28*	51	17
12	E-E	108	140	37	33	52	18
13	**Farmers in control villages**	**359**	**145**	**38**	**35**	**53**	**19**
14	**Missings**[1]	**147**	**137**	**34**	**33***	**50***	**20**
15	T	47	139	36	35	49	19
16	R	5	126	20	31	55	20
17	E	22	143	36	36	51	21
18	C	73	135	34	31	51	20

Table 7.26b. Difference of farmer test score between 2011 and 2012 in Anhui, China

		Delta of total score[5]	Delta of each component			
			Fertilizer[6]	Pest[6]	Cultivation[6]	Environment[6]
1	Total	19	7	6	3	3
2	**FFS graduate in treatment villages**[1]	**24****	**7**	**8*****	**4**	**5***
3	T-T	24	7	9	4	4
4	R-T[2]	37	6	13	13	6
5	E-T[2]	25	9	8	2	6
6	**Non-compliance farmers in treatment villages**[1]	**47*****	**16***	**9****	**10***	**12****
7	T-R[3]	30	11	6	8	5
8	R-R	5	2	5	-1	-1
9	E-R[3]	57**	20*	10	10	18*
10	**Exposed farmers in treatment villages**[1]	**10**	**4**	**4**	**2**	**0**
11	T-E[4]	13	3	6	6	-2
12	E-E	13	5	4	2	2
13	**Farmers in control villages**	**12**	**5**	**3**	**2**	**1**

Note: The code of the farmer type is explained in the text above.
1 t-test is conducted be referring to farmers in control villages (row 13).
2 t-test is conducted by referring to T-T (in row 3).
3 t-test is conducted by referring to R-R (in row 8).
4 t-test is conducted by referring to E-E (in row 12).
5 Full marks=400
6 Full marks=100

Table 7.27 shows the details of the knowledge test of greenhouse tomato farmers. Through the detailed comparison of knowledge scores between pre- and post-intervention surveys by group, we find that farmers in the treatment group get slightly higher knowledge scores as a whole and in the fertilizer test, and a more noticeable improvement in the environment protection test.

Compared with rice farmers, tomato farmers have higher knowledge scores. The main explanation is that the education level of the tomato farmers (8 years) is much higher than that of rice farmers (4.8 years). As mentioned earlier, the tomato farmers have higher incomes, and the higher education is presumably a contributing factor.

Table 7.27. Comparison of test scores between treatment and control groups of tomato farmers in Hebei, 2011–2013 China

	Group	Total test score	Test score about fertilizer	Test score about pesticide	Test score about cultivation	Test score about environment protection
Baseline	TT	58.98	65.98	59.21	59.07	35.75
	CC	58.65	64.28	59.76	58.25	31.92
	P-value	0.77	0.24	0.65	0.61	0.08
Endline	TT	65.63	69.95	67.65	62.82	36.79
	CC	65.35	69.66	67.66	62.99	33.37
	P-value	0.79	0.87	0.99	0.92	0.22

7.5. Robustness checks

The optimal range of fertilizer for rice use was proposed by soil scientists based on local experiments, then calibrated by local extension experts based on local soil type, seed varieties and cropping patterns. Most of this work was based on field communication. In the DID in distance analysis in Sections 7.1 and 7.2, we defined distance as the absolute distance from the whole optimum range of the 165–180 kg/ha. In this sub-section, we try three different definitions: distance from middle point of optimum range, distance from the lower bound of optimum range, and distance from the upper bound of optimum range to test the robustness of our results. The results of the three different regressions (Tables 7.28 through 7.30) are similar to those obtained using the whole optimum range.

Table 7.28. Regression of differences in distance from middle point of optimum range (rice)

	(1)	(2)
FFS Treatment	-16.176***	-14.495**
	(-3.39)	(-2.83)
Education		0.47
		(0.67)
Female		1.203
		(0.19)
Years farming rice		0.149
		(0.67)
Organic		7.122
		(0.63)
Own consumption		-8.072
		(-1.44)
Time on off-farm work		0.052
		(0.53)
Number of total plots		1.166
		(1.51)
Mobile phone use		11.094
		(1.9)
Work with other farmers		7.832
		(0.9)
Cost of pesticide		0.022
		(1.45)
Cost of fertilizer		-6.459*
		(-2.18)
_cons	8.516*	-4.291
	(2.53)	(-0.31)
F-test	11.51	2.81
Number of obs 711		653

Note: *Organic* is an indicator variable that is 1 if the farmer uses organic fertilizer; *own consumption* is an indicator variable if the farmer's household consumes the crop grown on the plot; * p<0.05, ** p<0.01

Table 7.29. Regression of differences in distance from lower bound of optimum range (rice)

	(1)	(2)
FFS Treatment	-14.920**	-13.352*
	(-3.10)	(-2.57)
Education		0.456
		(0.64)
Female		0.938
		(0.14)
Years farming rice		0.069
		(0.3)
Organic		6.065
		(0.54)
Own consumption		-9.422
		(-1.66)
Time on off-farm work		0.03
		(0.3)
Number of total plots		1.175
		(1.5)
Mobile phone use		9.651
		(1.63)
Work with other farmers		8.42
		(0.95)
Cost of pesticide		0.021
		(1.38)
Cost of fertilizer		-6.930*
		(-2.30)
_cons	4.704	-2.304
	(1.38)	(-0.16)
F-test	9.58	2.69
Number of obs	715	656

Note: *Organic* is an indicator variable that is 1 if the farmer uses organic fertilizer; *own consumption* is an indicator variable if the farmer's household consumes the crop grown on the plot; * p<0.05, ** p<0.01

Table 7.30. Regression of differences in distance from upper bound of optimum range (rice)

	(1)	(2)
FFS Treatment	-16.450***	-14.821**
	(-3.48)	(-2.92)
Education		0.457
		(0.66)
Female		1.343
		(0.21)
Years farming rice		0.187
		(0.85)
Organic		8.064
		(0.73)
Own consumption		-8.321
		(-1.50)
Time on off-farm work		0.043
		(0.43)
Number of total plots		1.149
		(1.5)
Mobile phone use		10.561
		(1.82)
Work with other farmers		7.596
		(0.88)
Cost of pesticide		0.018
		(1.24)
Cost of fertilizer		-5.874*
		(-1.98)
_cons	11.149***	-2.554
	(3.35)	(-0.18)
F-test	12.12	2.66
Number of obs	715	655

Note: *Organic* is an indicator variable that is 1 if the farmer uses organic fertilizer; *own consumption* is an indicator variable if the farmer's household consumes the crop grown on the plot; * p<0.05, ** p<0.01

7.6. Diffusion effects on exposed farmers

Farmer-to-farmer knowledge diffusion implies that untreated farmers in the same village as treated farmers may change their farming behaviour, presumably by coming into contact with treated farmers.

7.6.1 Comparison of the pure treatment and exposed groups

Table 7.31 compares the mean fertilizer usage for the pure treatment and pure exposed groups. In later sub-sections we compare results for other groups.

Table 7.31. Exposed effect of FFS on fertilizer use (rice)

	Sample	Nitrogen			Potassium		
		Baseline	Endline	Delta	Baseline	Endline	Delta
Control (C-C)	359	174	137	-37	35	43	8
Treatment(T-T)	356	180	148	-32	32	46	14
Exposed(E-E)[1]	108	178	131*	-47	36	40	4*

Note: * *p<0.05, ** p<0.01; 1 t-test is conducted by referring to T-T* (t-test is also conducted by referring to C-C, but there is no significant difference between E-E and *C-C*)

Endline nitrogen fertilizer use fell dramatically in *both* treatment group and exposed group, and the reduction is even higher in the exposed group. The decrease in fertilizer use in the exposed group could be due to knowledge diffusion from FFS graduates to exposed farmers, but we do not have clear evidence of this effect. Another reason may be the unexpected weather pattern in Anhui in the baseline survey year 2011 that we discussed earlier. The common weather could explain the reduction in fertilizer use in the exposed group as it does in the control group, making it difficult to identify a specific diffusion effect.

The application of potassium fertilizer increased post-FFS in all three groups; the treatment group shows the most pronounced increase. The increase of potassium fertilizer use in the exposed group is smaller than in the control group, which suggests there is no evidence of diffusion effects.

Table 7.32. Exposed effect of FFS on fertilizer use (tomato)

	Sample	Nitrogen			Potassium		
		Baseline	Endline	Delta	Baseline	Endline	Delta
Control (C-C)	206	501	488	-13	628	588	-40
Treatment(T-T)	193	368	482	114	456	591	134
Exposed(E-E)[1]	39	381	781**††	401**††	431	940**††	509**††

Note: * *p < 0.05, ** p < 0.01; t-tests refer to comparisons between T-T*(*) and *C-C* (†).

Based on Table 7.32, there is no significant difference in the baseline fertilizer use between the treatment group and the exposed group, which makes the exposed group a valid counterfactual (with the caveat of the much smaller sample size). The application of nitrogen and potassium fertilizer increased after intervention in both the treatment and exposed groups, but the exposed group shows a substantially larger increase, and changes we cannot explain and do not attribute to knowledge spillovers from the treated farmers.

Since fertilizer usage is highly heterogeneous among farmers, we break down N fertilizer usage by quintile. This is shown in Table 7.33.

Table 7.33. Comparison of N fertilizer usage by quintile in rice

Group	N		Mean	Quartile				
				0–20	20–40	40–60	60–80	80–100
				n=72	n=71	n=71	n=71	n=71
TT	356	nitrogen baseline	180	69	124	164	223	320
		nitrogen endline	148	114	124	142	152	206
		delta of nitrogen	-32	45	0	-22	-71	-114
		% change	-18%	65%	0	-13%	-32%	-36%
				n=22	n=22	n=21	n=22	n=21
EE	108	nitrogen baseline	178	78	136	172	216	292
		nitrogen endline	131	93	126	129	134	174
		delta of nitrogen	-47	15	-10	-43	-82	-118
		% change	-26%	19%	-7%	-25%	-38%	-40%
				n=72	n=72	n=72	n=72	n=71
CC	359	nitrogen baseline	174	75	128	167	213	287
		nitrogen endline	137	95	114	132	154	190
		delta of nitrogen	-37	20	-14	-35	-59	-97
		% change	-21%	27%	-11%	-21%	-28%	-34%

Nitrogen use increased in the lowest quintile in all the groups while nitrogen use fell for the other quintiles (except in the 20–40 per cent quintile in the treatment group, where it stayed the same). The reduction is the highest in the topmost quintile; however, the reduction in the other four quintiles for the exposed group is larger than in the treated or control group, while the increase in the first quintile for the exposed group is lower than in the other two groups. Table 7.34 shows potassium usage by quintile.

Table 7.34. Comparison of K fertilizer usage by quintile in rice

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
TT	356			n=114	n=29	n=76	n=76	n=61
		potassium baseline	32	0	13	28	46	87
		potassium endline	46	33	38	46	50	68
		delta	14	33	25	18	4	-19
		% change	44%	-	192%	64%	9%	-22%
EE	108			n=27	n=17	n=23	n=20	n=21
		potassium baseline	36	0	17	33	49	87
		potassium endline	40	27	30	37	70	43
		delta	4	27	13	4	21	-44
		% change	11%	-	76%	12%	43%	-51%
CC	359			n=95	n=52	n=74	n=75	n=63
		potassium baseline	35	0	17	31	48	93
		potassium endline	43	34	36	47	43	54
		delta	8	34	19	16	-5	-39
		% change	23%	-	112%	52%	-10%	-42%

Potassium use increased dramatically in the lowest quintile in all the groups. The use of potassium fell significantly in both groups in the highest quintile, but the observed reduction in the exposed group is much higher compared to the reduction in the other two groups. In addition, the increase in the first three quintiles (0–60 per cent) for the exposed group is lower than that in the control group.

In summary, the reduction in fertilizer usage of rice farmers in the highest quintiles is higher in the exposed group than in the control group (and surprisingly even higher than in the treatment group), which is suggestive but not strong evidence of diffusion effects.

Table 7.35. Comparison of N fertilizer usage by quintile in tomato

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
TT	193			n=39	n=39	n=38	n=39	n=38
		nitrogen baseline	368	108	218	291	428	803
		nitrogen endline	482	441	332	463	450	729
		delta of nitrogen	114	333	114	172	22	-74
		% change	31%	308%	52%	59%	5%	-9%
EE	39			n=8	n=8	n=8	n=8	n=7
		nitrogen baseline	381	186	251	315	429	771
		nitrogen endline	781	1180	643	871	649	533
		delta of nitrogen	400	994	392	556	220	-238
		% change	105%	534%	156%	177%	51%	-31%
CC	206			n=42	n=41	n=41	n=41	n=41
		nitrogen baseline	501	155	269	367	492	1233
		nitrogen endline	489	399	413	444	470	719
		delta of nitrogen	-12	244	144	77	-22	-514
		% change	-2%	157%	54%	21%	-4%	-42%

For tomato farmers (Table 7.35), the reduction of nitrogen fertilizer use is the highest in the top quintile; however, the reduction in the control and exposed groups is larger than in the treated group. In addition, the increase in the first three quintiles for the exposed group is higher than in the other two groups.

Potassium use increased dramatically in the first two quintiles in all the groups, especially the exposed group (Table 7.36). The use of potassium reduced in the topmost quintile in the treatment and control groups, but increased in the exposed group. As we mentioned above, the sample size of each bin in the exposed group is much smaller than the others.

Table 7.36. Comparison of K fertilizer usage by quintile in tomato

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
TT	193			n=39	n=39	n=38	n=39	n=38
		potassium baseline	456	123	252	361	499	1058
		potassium endline	591	398	509	571	578	904
		delta of potassium	135	275	257	210	79	-154
		% change	30%	224%	102%	58%	16%	-15%
EE	39			n=8	n=8	n=8	n=8	n=7
		nitrogen baseline	431	125	276	387	505	925
		nitrogen endline	940	1600	811	514	805	976
		delta of nitrogen	509	1475	535	127	300	51
		% change	118%	1180%	194%	33%	59%	6%
CC	206			n=42	n=41	n=41	n=41	n=41
		potassium baseline	628	170	315	436	602	1627
		potassium endline	588	524	464	431	701	821
		delta of potassium	-40	354	149	-5	99	-806
		% change	-6%	208%	47%	-1%	16%	-50%

7.6.2. DID in distance from the optimum

Tables 7.33 and 7.34 are consistent with a diffusion effect of FFS knowledge for rice farmers, but general changes in behaviour common to all farmers could explain these observations. As before, we run regressions for the differences in distance from the optimum between endline and baseline on participation in FFS, exposed farmers, and other unbalanced variables and farmer household controls (Table 7.37). The two regressions show that although participation in FFS is significantly and negatively associated with this difference in distance, the diffusion effect is not significant in the exposed group. In other words, unlike the corresponding regressions for treatment farmers, exposed farmers did not reduce the distance from the optimum significantly. And the regressions for tomatoes (Table 7.38) show increased use of fertilizers for tomato farmers (though with marginal significance) and no effect for treated farmers.

Table 7.37. Regression of differences in distance from optimum (rice)

	(1)	(2)
Exposed	-4.595	-8.65
	(-0.67)	(-1.20)
FFS	-15.731***	-15.187**
	(-3.35)	(-3.03)
Education		0.922
		(1.43)
Female		-2.722
		(-0.46)
Years farming rice		0.148
		(0.71)
Organic		13.13
		(1.3)
Own consumption		-12.379*
		(-2.40)
Time on off-farm work		0.022
		(0.24)
Number of total plots		0.853
		(1.22)
Mobile phone use		6.932
		(1.3)
Work with other farmers		9.222
		(1.18)
Cost of pesticide		0.012
		(0.85)
Cost of fertilizer		-5.843*
		(-2.12)
_cons	7.926*	3.389
	(2.4)	(0.27)
F-test	5.75	2.76
Number of obs	823	756

Note: *Organic* is an indicator variable that is 1 if the farmer uses organic fertilizer; *own consumption* is an indicator variable if the farmer's household consumes the crop grown on the plot; * p<0.05, ** p<0.01

Table 7.38. Regression of differences in distance from optimum (tomato)

	(1)	(2)
Exposed	106.595	107.438
	(1.87)	(1.85)
FFS	-0.797	5.698
	(-0.02)	(0.17)
Education		-4.771
		(-0.79)
Male		35.331
		(0.95)
Age		-1.515
		(-0.80)
Years farming vegetable		2.172
		(0.8)
Time on off-farm work		1.553
		(0.05)
Other training		0.488
		(0.9)
Constant	28.655	71.035
	(1.18)	(0.71)
F-tests	1.95	0.90
N	324	324

Note: *Organic* is an indicator variable that is 1 if the farmer uses organic fertilizer; *own consumption* is an indicator variable if the farmer's household consumes the crop grown on the plot; * $p < 0.05$, ** $p < 0.01$

7.7 Intent-to-treat analysis

Since the participation or non-participation of selected farmers in the FFS does not appear to have been driven purely by characteristics inherent to the programme as much as the way the design was implemented in the field, we have relied primarily on treated-on-treated and IV analysis to understand programme effectiveness. However, in this section, we also present an intent-to-treat analysis, comparing the treatment and control groups as randomized.

Table 7.39 Comparison of N fertilizer usage by quintile (rice)

Group	N		Mean	Quintile 0–20	20–40	40–60	60–80	80–100
				n=81	n=82	n=80	n=81	n=81
		nitrogen baseline[1]	180	69	123*	165	225**	317*
T	405	nitrogen endline[2]	148	116*	120	141	153	210
		delta of nitrogen[3]	-32	47*	-3	-24	-72	-107
		% change	-18%	68%	-2%	-15%	-32%	-34%
				n=72	n=72	n=72	n=72	n=71
		nitrogen baseline	174	75	128	167	213	287
C	359	nitrogen endline	137	95	114	132	154	190
		delta of nitrogen	-37	20	-14	-35	-59	-97
		% change	-21%	27%	-11%	-21%	-28%	-34%

Note: * *p < 0.05*, ** *p < 0.01*
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

In Table 7.39, as in Table 7.13, nitrogen use increased in the first quintile (0–20 per cent) in both control and treatment groups while nitrogen use reduced for the other quintiles.[12] The reduction is the highest in the top quintile compared to other quintiles in both treatment and control groups. However, the increase in the first quintile for the treatment group (68 per cent) is substantially higher in the treatment than the control group.

In Table 7.40, as in Table 7.14, nitrogen use increased in the first quintile (0–20 per cent) in both control and treatment groups, but the increase in the treatment group is significantly higher than in the control group. The reduction is the highest in the top quintile in both treatment and control groups, and actually higher for the control group (though relative to Table 7.14, there is more of a decrease for the treatment group). Similar trends can be found in Table 7.41, when Yon Qing is omitted (the decrease in the highest quintile in this case is more comparable to that in Table 7.15).

[12] Note that there is no difference in the control groups between the two tables since the CC and C groups are the same.

Table 7.40. Comparison of N fertilizer usage by quintile (tomato)

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
T	220			n=44	n=44	n=44	n=44	n=44
		nitrogen baseline[1]	444	106*	221**	302**	452**	1140
		nitrogen endline[2]	492	444	349	407	472	787
		delta of nitrogen[3]	48	338	128	105	20	-353
		% change	11%	319%**	58%	35%	4%	-31%
C	206			n=44	n=44	n=44	n=44	n=44
		nitrogen baseline	501	155	269	367	492	1233
		nitrogen endline	489	399	413	444	470	719
		delta of nitrogen	-12	244	144	77	-22	-514
		% change	-2%	157%	54%	21%	-4%	-42%

Note: * *p<0.05*, ** *p<0.01*
1 t-test is conducted be referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

Table 7.41. Comparison of N fertilizer usage by quintile (tomato) without Yong Qing county

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
T	181			n=37	n=36	n=36	n=36	n=36
		nitrogen baseline[1]	321	102**	206**	271**	380**	653
		nitrogen endline[2]	426	452	341	413	376	547
		delta of nitrogen[3]	105	350	135	142	-4	-106
		% change	33%	343%*	66%	52%	-1%	-16%
C	154			n=44	n=44	n=44	n=44	n=44
		nitrogen baseline	365	145	246	326	432	686
		nitrogen endline	415	422	359	394	424	478
		delta of nitrogen	50	277	113	68	-8	-208
		% change	14%	191%	46%	21%	-2%	-30%

Note: * *p<0.05*, ** *p<0.01*
1 t-test is conducted by referring to baseline nitrogen of corresponding quintile bin in control villages.
2 t-test is conducted by referring to endline nitrogen of corresponding quintile bin in control villages.
3 t-test is conducted by referring to delta of nitrogen of corresponding quintile bin in control villages.

8. Cost-effectiveness analysis

In this section, we report results from our cost-effectiveness analysis to inform future decisions about continuing or expanding the FFS programme in China.

8.1 Programme implementation costs

Most studies on rural education used the criterion of cost per farmer to assess the relative advantage of the project. We report total implementation costs and costs on a per famer basis (Ooi *et al*. 2005). Farmer field school expenditures consist of start-up (fixed) costs and operating (variable) costs. Fixed costs included the training of trainers (TOT) workshop and a subsequent motivation workshop; costs that do not scale directly with the number of villages targeted or farmers trained. Variable costs are those which the MoA incurred when running the FFS programme in the field, including technical assistance by the project management unit, travel, equipment and so on. We focus primarily on variable costs, because fixed costs could vary significantly in a future FFS rollout. For example, expanding the number of villages in Hebei and Anhui provinces might require little or no additional investment in training for MoA staff. Nevertheless, we report estimates of fixed costs wherever feasible.

Our data on cost was provided by the MOA, and reflects the best available estimate of programme costs. According to the MOA, the expense per rice FFS is approximately US$2,500, while the cost per tomato FFS is about US$3,300. These numbers exclude fixed costs. We randomly selected 28 rice villages and 18 tomato villages to receive the training and each village had one FFS. Hence, the total implementation cost is US$70,000 for rice and US$59,400 for tomatoes; the total variable costs to the MoA for the FFS programme we evaluated was US$129,400. On a per farmer basis, the average implementation cost was approximately US$70.

Fixed costs, including TOT workshops and motivation workshop, were approximately US$60,000. These costs are preliminary estimates, since we are awaiting additional cost information from the MoA.

Farmers incurred opportunity costs while participating in the training, which we have not included in the analysis.

8.2 Benefits associated with FFS expenditures

We can calculate the benefits farmers obtained from the FFS programme based on the impact analysis in Section 5. Impacts include a mix of public and private benefits, and not all impacts make sense when interpreted on a per dollar cost basis—especially those that are not easily monetized like farmer knowledge. To review, the expected benefits from FFS are:

Optimized use of fertilizer

For rice farmers, the FFS is associated with an approximately 15 point 'optimization' of fertilizer application (change in fertilizer use in kilograms per hectare in the direction of the optimal range). For farmers at the upper end of the distribution, this results in a net reduction of fertilizer, with associated public and private benefits. Aggregate N fertilizer reductions for rice and tomato farmers who reduced their use between baseline and endline were 56,171 kg/h and 70,854 kg/h, respectively. However, farmers at the lower end of the distribution increased fertilizer use, at personal cost and with some costs to the environment. Total increases were 17,938 kg/h and 100,341 kg/h for rice and tomato farmers, respectively.

For this reason, calculating the net social benefits is not feasible, although this will be the topic of future analysis. We can report that the variable cost associated with moving fertilizer closer to the optimum for rice farmers (based on average treatment effects) was US$~5 per kg per hectare.

Enhanced farmer knowledge and skills

Farmer field school graduates scored higher post-FFS compared to farmers in the control group. Specifically, rice farmers in the treatment group increased their knowledge score by 24 points, while those in the control group saw half the increase FFS graduates did (12 points) for a net gain of 12 points. On a per dollar basis, this equates to roughly US$6 per point gained, although we argue that this is not necessarily a useful way to look at costs because we cannot monetize knowledge gains in our experiment. In particular, some of the knowledge gained is likely to be general and benefit other farming (and perhaps even non-farming) activities the farmers might undertake.

For tomato farmers, the increase in knowledge is smaller and equivalent across treatment and control groups: about six points in both. The FFS programme provides no net increase in knowledge.

Enhanced environmental awareness

Enhanced environmental awareness is hard to measure and more difficult to monetize, but farmers in the FFS did become more aware of the importance of environmental protection. We divided the test for tomato farmers into sections, and for the environment protection section farmers achieved an approximately five-point increase compared to the control group.

Increased yield

Rice farmers saw an approximately 9 per cent increase in yield in both treatment and control groups. In contrast, tomato farmers who attended the FFS showed about an 18 per cent increase, while control group farmers increased by about 5 per cent.

Spillover effects

Positive spillover (or diffusion) effects would increase the value of the FFS per dollar spent if non-treated farmers in treated villages achieved indirect benefits from the FFS programme, e.g. through knowledge sharing with FFS farmers. However, we observed little evidence of positive spillovers, especially for tomato growers, and the beneficial effect on costs is likely minimal or non-existent.

9. Observations and policy recommendations

China is the largest fertilizer user in the world, and its chemical fertilizer use per hectare is also one of the highest in the world. The current rate of nitrogen fertilizer use for many farmers in China not only does not significantly improve crop yield, but also leads to serious food safety and environmental problems. The FFS is a mechanism being used in China and other countries to improve farmer knowledge and farming outcomes. This study used an RCT impact evaluation to estimate the causal effect of FFS programmes on environmental and economic outcomes. We find evidence that the FFS programme worked primarily for farmers who use less than the recommended amounts of fertilizer. We find evidence of reduced fertilizer application for high users only for rice growers, and overall programme impacts were only significant for rice farmers. Overall, results were inconclusive.

Reforming agricultural extension services is a priority in China's agricultural agenda. After three years of pilot FFS projects that disseminated knowledge to greenhouse vegetable farmers in Beijing, the MoA is considering using FFS as one of the core tools for China's agricultural extension programme. The effectiveness of FFS in reducing excess fertilizer use—and its associated environmental and social-economic impacts—is critical information that will be used by the MoA in decisions concerning scaling up its national FFS programme in the upcoming years. However, our results do not provide unambiguous evidence that could guide China's decision making on future FFS use. Here we outline a series of recommendations based on what we can draw from our results, despite an unclear signal about the benefits of FFS.

Observations

This evaluation was conducted in an institutional environment that led to inconsistent implementation outcomes, which provides insights into the potential and limitation of FFS scale-up. In our experience, the enthusiasm with which the FFS was embraced and the care with which the protocols were implemented varied significantly by region and extension service. Not all county officials or village heads were equally motivated to ensure that the FFS was properly implemented and had a chance to succeed. The recruitment of farmers, sometimes from the explicitly earmarked exposed or refused groups, in order to satisfy a minimum enrollment criterion is but one of the challenges we encountered. Parallel training sessions by fertilizer dealers who could have objectives antithetical to the FFS was another. If our surveys taught us that the farmer use of fertilizers was highly heterogeneous, our field experience taught us that the institutional environment was no less heterogeneous. To us these challenges seem symptomatic of a deep institutional heterogeneity that needs to be first addressed before issues of broad-based adoption and scaling up can even be addressed in a systematic way. It is in this spirit that the recommendations we give in this section are to be taken.

The evaluation results are inconclusive about overall FFS effectiveness, suggesting that policymakers should revisit plans to scale up FFS in China in the near future. A primary motivation for conducting this study was to guide decisions about whether to scale up FFS programmes throughout China to deal with excessive fertilizer use. Based on the results discussed above, we find some evidence—particularly for rice farmers—that FFS participation is significantly associated with balanced fertilizer use by moving farmers closer to the agronomic optimum. However, the changes in fertilizer use are not dramatic, particularly for reducing overuse, and the effects for tomato farmers are much weaker. While there is some evidence—again, for rice farmers—that the treatment increased farming knowledge and increased yields, control groups saw substantial changes in these outcomes without the FFS mechanism. Overall, while the treatment had some positive impacts, we do not believe the results unambiguously recommend broad-based use or scale-up of the FFS programme. However, it is useful to note that the FFS was implemented as a package, not just as a programme to influence fertilizer usage and there may be other aspects of the programme (for instance, imparting knowledge about farming practices as a whole, which is reflected in improved test scores) that could be beneficial.

Recommendations

It is not obvious that the FFS programme is cost-effective, and the MoA should proceed cautiously when considering how to change or expand the programme and look for ways to reduce costs. FFS is a relatively expensive approach to improving farm practices, and cost-effectiveness has always been a big concern for extension agents and policymakers. Our analysis found that an FFS programme costs between US$2,500 and US$3,300 per village (~US$100–130 per farmer), depending on the crop.[13] A full cost-benefit analysis is necessary to determine whether social benefits outweighed costs to the MoA, but that analysis is beyond the scope of this study. However, given the ambiguous impacts on fertilizer, farmer knowledge and crop yields, we cannot conclude that the FFS programme we evaluated *was* cost-effective. The MoA should consider carefully the value of different outcomes (including private benefits to farmers) when considering scaling up the FFS programme. In

[13] These amounts include only variable costs, not fixed costs; including fixed costs will only increase the average cost per village in many cases further decreasing cost-effectiveness.

addition, the MoA should identify ways to reduce programme costs or spread those costs over a wider beneficiary pool. For example, one of the most expensive components of the FFS programme was the training workshop. If these workshops could be scaled up without loss of quality, it would reduce the cost per agent trained. In addition, although we calculated costs per farmer for this FFS programme, activities like extension agent training will presumably have benefits before this particular FFS, which increases cost-effectiveness.

FFS training should be designed and rolled out based on local needs, including crop-specific considerations. Our evaluation results were heterogeneous across crops and across the distribution of fertilizer use within crops, and this should inform future FFS design. For example, fertilizer impacts were largest for rice farmers, but these farmers have lower fertilizer use on average. In addition, observed farmer characteristics vary by crop, which could affect the FFS impact. For example, rice farmers saw the largest improvements in farmer knowledge, but these farmers have lower educational attainment than tomato farmers, which could be driving the result. Future FFS for rice could focus more on increasing general knowledge and yield-improving techniques, while tomato FFS could be designed to leverage farmer knowledge and place greater emphasis on fertilizer use (including reduction). Any existing agricultural census (conducted for gross domestic product calculations) could help identify these specific target groups. The MoA would also benefit from conducting additional evaluations of the FFS for other crops (e.g. other GHV) to explore the crop-specific benefits of the FFS.

FFS implementation quality was heterogeneous, and the MoA should assess ways to improve implementation and quality control if they decide to continue to use or expand the FFS programme. As in most high-quality impact evaluations, we worked closely with the implementer (MoA) to ensure that the programme was carried out as closely to the experimental design as possible. Nevertheless, the extension staff in different provinces and counties conducted the FFS with varying levels of motivation and dedication. Consequently, the results from our experiment have high external validity if no major changes are made to FFS implementation. In other words, we believe the results could have been stronger had the programme implementation been carried out under more stringent 'laboratory' settings. If the MoA decides to expand FFS use, we recommend they focus on improving the quality of FFS implementation where feasible. Our data collection was not designed to specifically identify factors that drive extension agent motivation, but the MoA could explore this issue in greater detail through focus groups or constructive discussions with agents. Obvious considerations are compensation and incentive schemes for agents and management structure and overseeing at county offices.

The MoA should acknowledge constraints that are external to the FFS programme and try to adapt to those when possible. Two challenges we encountered that we believe reduced—or at least obscured—the effects of the intervention were off-farm work behaviour and training provided by nonextension workers. Many farmers in rice counties spent significant time working off the farm, which led to high dropout rates and reduced farming effort. The MoA cannot affect the benefits of off-farm work, but it could adapt the FFS to: (a) emphasizing the private benefits of training (i.e. increased yields); or (b) targeting other household members with the FFS programme—those who aren't absent for long periods of time. The other challenge was with training provided by agriculture firms or others outside the extension system. These programmes could either supplement or counteract FFS training, and the MoA should work to: (a) better understand what information these programmes convey; and (b) if and when the non-MoA training is complementary, leverage that training or reduce FFS effort in areas where farmers have access to other programmes that provide similar information.

Appendix A: Sample design

General notes

- Clustered, blocked randomized control trial
  - FFS programme delivered at village level = cluster level
  - T & C villages to be matched prior to group assignment
- Power calculations and sampling scheme for Hebei (greenhouse tomatoes) and Anhui (rice) provinces are separate and different
  - need to credibly claim pre-post changes separately for each crop
  - FFS programme is tailored to each crop, therefore treatment will differ
- TREATMENT VILLAGES
  - FFS target = 25 FFS participants, of whom 18 will be identified from the beginning, followed and surveyed in treatment villages
  - therefore in smaller villages where the number of households < 37 (25 FFS + 12 exposed), we have 2 options:
    - no exposed group
    - exposed group comprising other GHV farmers

Hebei (greenhouse tomatoes): Gao Cheng (10 villages), Yong Qing (10 villages), and Rao Yang (16–18 villages)

- 36 villages for 20% pre-post change
- survey target: 15 FFS, 10 exposed farmers and X non-compliers per treatment village; 15 farmers and X non-compliers per control village = **720 + 36X farmers**
  - to account for 15% attrition, recruit: 30 (18T + 12E) and X non-complying farmers in treatment villages and 18 farmers + X non-compliers in control villages = **864 + 36X farmers**
- sampling scheme
  - 3 counties chosen based on 1) total sowing area for greenhouse tomatoes and 2) willingness of county government to participate in study
  - select all villages available: 36
  - match and assign T & C villages after some village data collection

Anhui (mid-season rice): Tian Chang and Ju Cao counties

- 56 villages for 15% pre-post change
- target: 15 FFS, 10 exposed farmers, and X non-compliers per treatment village; 15 farmers and X non-compliers per control village = **1,120 farmers + 36X farmers**
  - to account for 15% attrition, recruit: 30 (18T + 12E) and X non-complying farmers in treatment villages and 18 farmers + X non-compliers in control villages = **1,344 + 56X farmers**
- sampling scheme

- 2 counties chosen based on 1) total sowing area for rice and 2) willingness of county government to participate in study
  - within each county, eliminate townships in which FFS is not possible for various reasons (see decision log), then sample to achieve geographic/terrain representativeness as cropping data is found to be inaccurate (e.g. divide county into 4 quadrants, and sample 1 township from each quadrant)
  - then randomly select 7 villages from each township
  - because rice villages are large in distance, randomly sample natural village from those in which number of rice households >= 37 (25 FFS + 12 exposed) + 60% refusal rate (estimate from fieldwork) = 60
  - match and assign T & C villages after some village data collection at both administrative and natural village levels

Farmer recruitment process

- eligibility criterion: must grow greenhouse tomatoes/rice this SPRING season (baseline) AND next SPRING season
- TREATMENT group in treatment villages
  - survey enumerators will screen and select households in the sample that meet the eligibility criterion until they reach 18 eligible households (round 1)
  - once a household has been determined to be eligible, survey enumerators will personally invite farmers to participate in the FFS programme by describing the nature of the programme, terms and conditions
  - farmers will have one full day to decide whether or not they would like to participate, after which survey enumerators will ask both refusing and accepting households to participate in baseline survey
  - all round 1 farmers will be surveyed, regardless of accepting or declining
  - in round 2 and thereafter, additional farmers will be invited equal to the number of declining farmers in the previous; however only accepting farmers from round 2 onward will be surveyed (see diagram below)
  - after we reach 18 target survey households, we will extend 7 more invitations to other households to fill the minimum FFS quota of 25

Example invitation process for first 18 households

10 accept → survey

18

8 decline → survey

invite 8 more

6 accept (survey) 2 decline (do not survey)

invite 6 more

all 6 accept (survey) 0 decline

- total invited: 32 households
- total declined: 10 households
- total declined and surveyed: 8 households
- true declining rate = 10/32 = 31.25%
  - surveyed declining rate = 8/32 = 25% (can use survey weights to upweight data)

- EXPOSED group in treatment villages
  - survey enumerators will screen and select households in the second sample that meet the eligibility criterion until they reach 12 eligible households

- once a household has been determined to be eligible, survey enumerators will ask whether it would be interested in participating in a survey
    - willing households will be interviewed

- CONTROL group in control villages
    - survey enumerators will screen and select households in the sample that meet the eligibility criterion until they reach 18 eligible households
    - once a household has been determined to be eligible, survey enumerators will ask whether it would be interested in participating in a survey
    - willing households will be interviewed

Appendix B: Survey instruments

Survey questionnaires and face-to-face interviews are the primary data collection tools. We designed the questionnaires to ask questions about regional-, institutional-, and household-level characteristics and outcomes.

Baseline and endline survey instruments are attached as separate documents.

Appendix C: Sample size and power calculations

Based on existing publications (Chen *et al*. 2006; Cui 2005; He *et al.* 2009; Ju *et al.* 2009; Zhu and Chen 2002), second-hand statistics (viz.,NDRC. 2010), previous fieldwork conducted by the Center for Chinese Agricultural Policy (CCAP) (Hu *et al.* 2007), and personal communication with local researchers, we obtained and calibrated means and standard deviations for nitrogen fertilizer usage rate for both types of crops: rice and GHVs. The two crops yielded similar standardized effect sizes, and we adopted the slightly more conservative one (rice) and conducted power calculations for a range of minimum desired effect sizes, from 10 to 20 per cent. In our power calculations, we have also allowed for various correlations between farmer outcome variables, ranging from 0.05 to 0.20. In the proposal we set *n*, or the number of farmers per cluster (village), at 20 treatment and 20 control (in FFS treatment villages, we also include 20 additional farmers who are in the exposed group), but it is feasible to use somewhat smaller group sizes, as the number of individuals within a cluster does not have as large an effect on overall power as the number of clusters itself.

We anticipate an approximately 3 per cent attrition rate in the survey sample, based on previous CCAP fieldwork in surveying farmers in which the authors found that farmers tended to drop out of similar agricultural studies at a rate of no more than 3 per cent (Huang *et al*. 2010).

We also predict a drop-out rate of no more than 10 per cent for FFS treatment group farmers based on the following reasons. First, the farmers will receive invitation letters before they make their decisions to participate in the FFS programme, thereby ensuring their interest in the programme. Second, the FFS programme is a new and participatory process that has been shown by previous pilot studies to attract farmers to participate in China. And third, farmers understand that knowledge of pest management and fertilizer use is highly related to yield outcomes and income.

Given the above, the choice of 20 farmers per group conservatively incorporates the range of attrition rates that we believe we will encounter over the duration of the study. Even if we were to lose five farmers per group (equivalent to an attrition rate of 25 per cent), we would retain enough power to detect the desired effects (see line highlighted in green in Table C.1). We believe this consideration is important for possible contingencies that may arise.

Table C.1 shows the power calculation parameters that were used in our power calculations. The power is set at 80 per cent and the significance level at 0.05. The top half of the table presents a conservative benchmark against which we varied each of the parameters, which are in turn highlighted in bold font. The bottom half of the table presents a slightly less conservative benchmark, varying the same parameters. The lines highlighted in grey indicate the parameters that allow us to remain within our survey budget, and the line highlighted in yellow indicates the design with the most conservative parameters that still lies within the survey budget. This is the sample design that we adopt. It assumes a 15 per cent minimum detectable change in nitrogen fertilizer usage before and after the FFS intervention. We allow intra-cluster correlation to reach a correlation of 0.10, and we conservatively assume that no other covariates will have any additional explanatory power. This sample design calls for 52 villages (26 treatment and 26 control) and 1,560 farmers. We conservatively account for attrition with the figure of 20 farmers per group.

Table C.1. Power calculations

	delta	rho	n	covariate	N	total sample size	# FFS
	alpha = 0.05, power = 80% Parameters						
conservative benchmark	**0.206**	0.20	20	0.00	182	5,460	91
	0.310	0.20	20	0.00	81	2,430	41
	0.413	0.20	20	0.00	48	1,440	24
	0.206	**0.15**	20	0.00	145	4,350	73
	0.206	**0.10**	20	0.00	111	3,330	56
	0.206	**0.05**	20	0.00	75	2,250	38
	0.206	0.20	**15**	0.00	191	4,298	96
	0.206	0.20	**10**	0.00	210	3,150	105
	0.206	0.20	20	**0.10**	166	4,980	83
	0.206	0.20	20	**0.15**	159	4,770	80
	0.206	0.20	20	**0.20**	151	4,530	76
less conservative benchmark	**0.310**	0.10	20	0.10	48	1,440	24
	0.413	0.10	20	0.10	28	840	14
	0.206	0.10	20	0.10	104	3,120	52
	0.310	**0.15**	20	0.10	61	1,830	31
	0.310	**0.20**	20	0.10	75	2,250	38
	0.310	**0.05**	20	0.10	34	1,020	17
	0.310	0.10	**15**	0.10	52	1,170	26
	0.310	0.10	**10**	0.10	62	930	31
	0.310	0.10	20	**0.15**	46	1,380	23
	0.310	0.10	20	**0.20**	44	1,320	22
	0.310	0.10	20	**0.00**	52	1,560	26

Appendix D: Descriptive statistics, univariate, and bivariate tabulations of main variables of interest

Summary of fertilizer use by group (rice)

	Sample	Nitrogen			Potassium		
		Baseline	Endline	Delta	Baseline	Endline	Delta
Treatment (T-T)	356	180	147	-32	32	46	14
Control (C-C)	359	174	137	-37	35	43	8
Difference b/w T-T and C-C		6	10	5	-3	3	6*
p-value		0.33	0.06	0.45	0.2	0.3	0.05

Note: * *p<0.05, ** p<0.01, *** p<0.001*

Summary of fertilizer use by group (tomato)

	N	Nitrogen			Potassium		
		Baseline	Endline	Delta	Baseline	Endline	Delta
Treatment (T-T)	193	368	482	114	456	591	134
Control (C-C)	206	501	488	-13	628	588	-40
Difference b/w T-T and C-C		-133*	-6	127**	-171*	3	174**
p-value		0.03	0.86	0.01	0.03	0.96	0.01

Note: * *p<0.05, ** p<0.01*

Appendix E: Analytical tables and results tables including econometric model specification and tables showing balance tests and results with standard errors/significance levels

We compared average characteristics of households from FFS with those of non-FFS villages, in terms of demographic characteristics, times of nutrient (fertilizer) and pesticide application, amount of nutrient and pesticide input, off-farm employment time, experience of rice agricultural skills training in the past three years, the number of total plots, the size of the biggest plot, cost of fertilizer and pesticides, measures of social network, and so on. Table E.1 shows that the equality in means between treatment group and exposed group cannot be rejected for almost all but two characteristics. Most (all but four) of the characteristics are equal in means between treatment group and control group. For tomatoes (Table E.2), equality of means again cannot be rejected only for a couple of characteristics. In other words, our randomization seems to have worked well to produce a balanced sample.

Table E.1. Balance table for rice

	Treatment group A1	Exposed group C	Test of means (treatment/ exposure)	Control group D	Test of means (treatment/ control)
Number of observation	450	247	-	432	-
Knowledge score of rice production (full mark 100)	35.7	35.6	0.84	36.6	0.15
Yield (kg/ha)	7434	7315	0.53	7245	0.31
Times of nutrient application	2	2.1	0.41	2.1	0.75
Total nutrient input (kg/ha)	324	334	0.61	314	0.56
N fertilizer use (kg/ha)	235	242	0.62	224	0.36
Times of pesticides application	2.8	2.8	0.86	2.9	0.11
Amount of pesticide use (kg/ha)	19.7	20.2	0.76	19.4	0.79
Sex (fraction of male)	0.53	0.55	0.54	0.58	0.12
Age	54	53	0.46	53	0.35
Education (years)	3.9	4.9	0.00	4.4	0.05
Experience of rice farming for the primary labor (years)	31.5	31.2	0.75	30.6	0.29
Fraction of farmers received advice from extension people in the rice production in 2011	0.27	0.23	0.29	0.23	0.13
Fraction of farmers received advice from agro-chemical sellers in rice production in 2011	0.73	0.76	0.35	0.75	0.48
Fraction of participated in rice agricultural skills training in the past 3 years	0.07	0.09	0.35	0.09	0.29
The number of total plots	6	6	0.45	7	0.00
The size of the largest plot growing middle-season long-grained rice (ha)	40	36	0.35	35	0.18
Cost of fertilizer (yuan/kg)	1.9	1.9	0.69	1.9	0.64
Cost of pesticides (yuan/kg)	71	79.8	0.54	106.7	0.00

The time engaged in off-farm employment (percentage)	0.59	0.59	0.83	0.57	0.14
Individual social network					
Fraction of mobile phone use	0.59	0.56	0.36	0.66	0.02
Fraction of having internet access at home	0.11	0.14	0.29	0.07	0.05
Fraction of farmers work together with other farmers	0.07	0.13	0.03	0.11	0.04

Table E.2. Balance table for tomato

	Treatment group	Exposed group	Test of mean (treatment/exposure)	Control group	Test of mean (treatment/control)
Number of observation	235	67		217	
Knowledge score of tomato production (full mark 100)	59	56	0.07	59	0.66
Yield (kg/ha)	71323	77829	0.31	69780	0.73
N fertilizer use (kg/ha)	327	340	0.63	353	0.24
Amount of pesticide use (kg/ha)	18.2	20.8	0.42	16.2	0.36
Sex (fraction of male)	0.6	0.7	0.18	0.8	0.02
Age	44.5	47.8	0.01	43.0	0.10
Education (years)	8	8	0.18	8	0.48
Experience of vegetable farming for the primary labor (years)	12	14	0.07	12	0.79
Fraction of participated in vegetable agricultural skills training in the past 3 years	0.34	0.42	0.29	0.45	0.02
The size of the largest plot growing tomato (ha)	0.09	0.07	0	0.09	0.56
Time work off farm (%)	9.6	7.8	0.57	11.1	0.57

Appendix F: Supplemental analysis to Section 5

In this section, we compare treatment + newly transferred treatment group (T-T + E-T) and control group (C-C). The inclusiveness of the treatment group is in between the inclusiveness of the groups presented in Sections 5.1 and 5.2 of the main text. The R-T individuals are excluded to reduce selection bias because it is unclear why group R participants, those who initially refused to participate in FFS, would agree to do so in the second round. Perhaps such individuals were pressured by extension staff or local officials, more so than the pressure exerted on the E group to attend. Because the E-T farmers are randomly recruited in the second round, in this step we evaluate the effect of FFS on fertilizer use under the RCT implementation scenario by treating both T-T and E-T as FFS graduates.

F.1. Difference in means

We compare the differences in mean fertilizer use between groups T-T + E-T (the treated by implementation) and group C-C (the control) in Table G.1.

Table F.1. Comparison of means for rice: (TT+ET) v CC

	Treatment (T-T + E-T)	Control (C-C)	Difference	p-value
Nitrogen Baseline	186	174	12*	0.04
Nitrogen Endline	149	137	12*	0.03
Delta in nitrogen	-37	-37	0	0.96
Potassium Baseline	33	35	-2	0.47
Potassium Endline	47	43	4	0.09
Delta in potassium	14	8	6*	0.04

Note: * *p<0.05, *** *p<0.01*

Unlike Table F.1, there is a significant difference in the baseline N fertilizer application between TT and CC, which means that the control group might not be a good counterfactual for the treatment group. The reduction of nitrogen application is the same in the two groups.

The application of potassium increased in both treatment and control groups after intervention but farmers in the treatment group used more K fertilizer than those in the control group.

Table F.2. Comparison of means for tomato: (TT+ET) v CC

	N	Nitrogen			Potassium		
		Baseline	Endline	Delta	Baseline	Endline	Delta
Treatment (T-T)	216	369	500	131	467	603	136
Control (C-C)	206	501	488	-13	628	588	-40
Difference b/w T-T and C-C		-133*	11	144**	-160	15	176**
p-value		0.02	0.77	0.001	0.03	0.76	0.01

As in Table F.2, the nitrogen fertilizer use of the treatment group in the baseline year is significantly smaller than that in the control group and the increase in the nitrogen fertilizer use of the treatment group is remarkable compared to the decrease in the control group. K fertilizer figures present similar changes in that the endline year K fertilizer use greatly increased in the treatment group while it slightly decreased in the control group.

We further break down fertilizer use by using quintile to examine the heterogeneous effects in Table F.3–Table F.6.

Table F.3. Comparison of N fertilizer usage by quintile (rice)

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
TT	460			n=92	n=92	n=92	n=92	n=92
		nitrogen baseline	185	72	130	172	232	323
		nitrogen endline	149	109	127	142	159	208
		delta of nitrogen	-36	32	-3	-30	-73	-115
		% change	-19%	44%	-2%	-17%	-31%	-36%
CC	359			n=72	n=72	n=72	n=72	n=71
		nitrogen baseline	174	75	128	167	213	287
		nitrogen endline	137	95	114	132	154	190
		delta of nitrogen	-37	20	-14	-35	-59	-97
		% change	-21%	27%	-11%	-21%	-28%	-34%

Nitrogen use increased in the first quintile (0–20%) in both control and treatment groups while nitrogen use reduced for the other quintiles. The reduction is the highest in the top quintile as compared to others in both treatment and control groups. However, the increase in the first quintile and decrease in the top two quintiles of the treatment group are higher than those in the control group. These results mirror those in Table F.3.

Table F.4. Comparison of potassium by quintile (rice)

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
				n=136	n=48	n=102	n=93	n=81
TT	460	potassium baseline	32	0	15	30	48	88
		potassium endline	46	33	40	44	55	71
		delta	14	33	25	14	7	-17
		% change	44%	-	167%	47%	15%	-19%
				n=95	n=52	n=74	n=75	n=63
CC	359	potassium baseline	35	0	17	31	48	93
		potassium endline	43	34	36	47	43	54
		delta	8	34	19	16	-5	-39
		% change	23%	-	112%	52%	-10%	-42%

Table F.4 shows K fertilizer usage by quintile. Potassium use increased dramatically in the first quintile (0–20%) in both control and treatment groups. The use of potassium reduced significantly in both the groups in the top quintile 80–100 per cent, but reduction in the control group is much higher as compared to the treatment group. Overall, this implies that farmers in the treatment group are using more K fertilizer than those in the control group, mirroring results in Table F.4.

Table F.5. Comparison of N fertilizer usage by quintile (tomato)

Group	N		Mean	Quintile				
				0–20	20–-40	40–-60	60–80	80–100
				n=44	n=43	n=43	n=43	n=43
TT	216	nitrogen baseline	369	110	217	294	428	801
		nitrogen endline	500	443	346	454	487	772
		delta of nitrogen	131	333	129	160	59	-29
		% change	36%	303%	59%	54%	14%	-4%
				n=42	n=41	n=41	n=41	n=41
CC	206	nitrogen baseline	501	155	269	367	492	1233
		nitrogen endline	489	399	413	444	470	719
		delta of nitrogen	-12	244	144	77	-22	-514
		% change	-2%	157%	54%	21%	-4%	-42%

Table F.6. Comparison of K fertilizer usage by quintile (tomato)

Group	N		Mean	Quintile				
				0–20	20–40	40–60	60–80	80–100
TT	216			n=44	n=43	n=43	n=43	n=43
		potassium baseline	467	129	257	370	511	1077
		potassium endline	603	399	523	570	572	956
		delta of potassium	136	270	266	200	61	-121
		% change	29%	209%	104%	54%	12%	-11%
CC	206			n=42	n=41	n=41	n=41	n=41
		potassium baseline	628	170	315	436	602	1627
		potassium endline	588	524	464	431	701	821
		delta of potassium	-40	354	149	-5	99	-806
		% change	-6%	208%	47%	-1%	16%	-50%

In Table F.5, nitrogen use increased in the first three quintiles (0–60%) in both control and treatment groups, but the increase in the treatment group is significantly higher than that in the control group. The reduction is the highest in the top quintile as compared to others in both treatment and control group. But the reduction in the top quintile of the control group is substantially higher than that in the treatment group. A similar trend can be found in the Table F.5.

F.2. DID in distances from optimal use

As in Table F.7, we examine if participation in FFS reduces the distance from the optimum for N fertilizer usage for the TT+ET group relative to control. The first two columns of Table F.7 show this is indeed the case.

Table F.7: Regression of differences in distance from optimum

	(1)	(2)
treatment	-14.79**	-15.048**
	(-3.30)	(-3.35)
education		0.761
		(1.23)
female		-1.897
		(-0.33)
years farming rice		0.146
		(0.75)
organic		-4.767
		(-0.48)
own consumption		9.577*
		(1.97)
Constant	7.926*	-14.995
	-2.36	(-1.19)
F-test	10.87	3.09

Note: *Organic* is an indicator variable that is 1 if the farmer uses organic fertilizer; *own consumption* is an indicator variable if the farmer's household consumes the crop grown on the plot; * p<0.05, ** p<0.01

References

Anhui Agriculture Information, Introduction of Agricuture, as of March 11, 2014, http://218.22.17.153/Sites/MainSite/List_2_2097.html

Economist, 2009. Climate change after Copenhagen: China's thing about numbers. 30 December.

Chen, X, Zhang, F, Römheld, V, Horlacher, D, Schulz, R, Böning-Zilkens, M, Wang, P, and Claupein, W, 2006. Synchronizing N supply from soil and fertilizer and N demand of winter wheat by an improved Nmin method. Nutrient Cycling in Agroecosystems, 74(2), pp. 91–98.

Cui , Z, 2005. Optimization of the nitrogen fertilizer management for a winter wheat-summer maize rotation system in the North China Plain from field to regional scale. PhD thesis. China Agricultural University, Beijing.

Gao, Q, 2008. Developing trends and theoretical system of agricultural extension. *Ancient and Modern Agriculture*. 4, pp. 17-23(in Chinese).

He, F, Jiang, R, Chen, Q, Zhang, F, and Su, F, 2009. Nitrous oxide emissions from an intensively managed greenhouse vegetable cropping system in northern China. *Environmental Pollution*, 157, pp. 1666–72.

Hebei Provincial Department of Agriculture, Introduction of Hebei, as of 11 March 2014, http://www.heagri.gov.cn/hbagri/nygk.jsp

Hu, R, Cao, J, Huang, J, Peng, S, Huang, J, Zhong, X, Zou, Y, Yang, J, and Buresh, R, 2007. Farmer participatory testing of standard and modified site-specific nitrogen management for irrigated rice in China. *Agricultural Systems*, 74(2), pp. 331-340.

Hu, R, Yang, Z, Kelly, P, and Huang, J, 2009. Agricultural extension system reform and agent time allocation in China. *China Economic Review*, 20(2), pp. 303–15.

Huang, J, Hu, R, Cao, J, and Rozelle, S, 2008. Training programs and in-the-field guidance to reduce China’s overuse of fertilizer without hurting profitability. *Journal of Soil and Water Conservation*, 63(5).

Huang, J, Jia, X, Xiang, C, Hu, R, and Hou, L, 2010. Farmers’ adoption of nitrogen management practice of upland summer maize in north China: an experimental design paper presented at the 117th EAAE Seminar Climate Change, Food Security and Resilience of Food and Agricultural Systems in Developing Countries: Mitigation and Adaptation Options, 25–27 November 2010.

Ju, X, Xing, G, Chen, X, Zhang, S, Zhang, L, Liu, X, Cui, Z, Yin, B, Christie, P, Zhu, Z and Zhang, F, 2009. Reducing environmental risk by improving N management in intensive Chinese agricultural systems. *Proceedings of the National Academy of Sciences of the United States of America*, 106(9), pp. 3041–46.

Mangan, J, and Mangan, MS, 1997. A comparison of two Ipm training strategies in China: the importance of concepts of the rice ecosystem for sustainable insect pest management. *Agriculture and Human Values*, 15, pp. 209–21.

National Development and Reform Commission in China, 2010. Compilation of information on cost and benefit of agricultural products. Beijing: China Statistics Press.

Ooi, PAC, Praneetvatakul, S, Waibel, H, and Walter-Echols, G, 2005. The impact of the FAO-EU IPM program for cotton in Asia. Pesticide Policy Project, Universität Hannover, Germany, Special Issue, Publication Series 9: 139.

PEW Climate Center, 2007. Climate change mitigation measures in the People's Republic of China. Available online: http://www.pewclimate.org/docUploads/International%20Brief%20-%20China.pdf

Quizon, J, Feder, G, and Murgai, R, 2001. Fiscal sustainability of agricultural extension: the case of the farmer field school approach. *Journal of International Agricultural and Extension Education*, 8, pp. 13–24.

Van den Berg, H, and Jiggins, J, 2007. Investing in farmers: the impacts of farmer field schools in relation to integrated pest management. *World Development*, 35(4), pp. 663–86.

Xing, GX, and Zhu, ZL, 2000. An assessment of N loss from agricultural fields to the environment in China. *Nutrient Cycling in Agroecosystems*, Volume 1.

Yang, P, Liu, W, Shan, X, Li, P, Zhou, J, Lu, J, and Li, Y, 2008. Effects of training on acquisition of pest management knowledge and skills by small vegetable farmers. *Crop Protection*, 27(12), pp. 1504–110.

Zhu, ZL, and Chen, DL, 2002. Nitrogen fertilizer use in China: contributions to food production, impacts on the environment and best management strategies. *Nutrient Cycling in Agroecosystems*, 63(2–3).

Publications in the 3ie Impact Evaluation Report Series

The following reports are available from
http://www.3ieimpact.org/evidence-hub/impact-evaluations-repository

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, *3ie Impact Evaluation Report 1*. Martinez, S, Naudeau, S and Pereira, V (2012)

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, *3ie Impact Evaluation Report 2*. Burwen, J and Levine, DI (2012)

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, *3ie Impact Evaluation Report 3*. Casey, K, Glennerster, R and Miguel, E (2013)

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A (2013)

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, *3ie Impact Evaluation Report 5.* Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M (2014)

The impact of daycare on maternal labour supply and child development in Mexico, *3ie Impact Evaluation Report 6*. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E (2014)

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, *3ie Impact Evaluation Report 7*. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P (2013)

Paying for performance in China's battle against anaemia, *3ie Impact Evaluation Report 8*. Zhang, L, Rozelle, S and Shi, Y (2013)

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, *3ie Impact Evaluation Report 9*. Ashraf, N, Bandiera, O and Jack, K (2013)

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N (2013)

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B (2014)

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R (2014)

Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R (2014)

Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14. Casaburi, L, Glennerster, R, Suri, T and Kamara, S (2014)

A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V (2014)

The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16. Banerji, R, Berry, J and Shortland, M (2014)

Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua, *3ie Impact Evaluation Report 17*. Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME (2014)

Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18. Brooker, S and Halliday, K (2015)

A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19. Cai, J, de Janvry, A and Sadoulet, E (2014)

Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme, *3ie Impact Evaluation Report 20*. Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P (2015)

Shelter from the storm: upgrading housing infrastructure in Latin American slums, *3ie Impact Evaluation Report 21*. Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R (2015)

Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G (2015)

A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana, 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M (2015)

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a *focus on gender impacts, 3ie Impact Evaluation Report 23*. Gine, X, Patel, S, Cuellar-Martinez, C, McCoy, S and Lauren, R (2015)

Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G (2015)

In China, a major agricultural challenge is the inefficient use of fertiliser and the environmental effects associated with overuse. The Chinese Ministry of Agriculture is trying to address this problem by instituting farmer field schools (FFS). Existing studies suggest that insufficient farmer knowledge and information about the effects of excessive fertiliser use is one reason for inefficient rates of nitrogen fertiliser application in China. This study evaluates the effectiveness of FFS training in reducing fertiliser use for rice and tomato crops in two provinces. Overall, the study concludes that the FFS programme improved the optimal use of fertiliser for rice farmers, but had insignificant effects for tomato growers.

Impact Evaluation Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

www.3ieimpact.org